A NEW NONPARAMETRIC TEST FOR INDEPENDENCE BETWEEN TWO SETS OF VARIATES

By

PETER WILLIAM GIESER

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1993

To my parents

and

the memory of William Trust

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to Dr. Ronald Randles, without whom this work would never have been completed. His constant encouragement and optimism were always sources of energy from which I could draw when things seemed bleakest. I would also like to thank the members of my supervisory committee, as well as all the faculty I have had a chance to get to know in the short time I have been at the University of Florida. They will never realize the enormous impact that their collective experience and knowledge has made on me. Special thanks go to Jane Pendergast for going beyond the call of duty and being willing to help me in ways not even related to statistics. To the many students whom I have met and become friends with, I wish to acknowledge my pleasure in having had the privilege of knowing them. I would especially like to thank Dan Bowling, who is probably one of the few people who could have put up with me for so long. I consider him among the best friends I have ever had. I am also indebted to Dr. James Kepner, who provided motivation via his excitement about statistics and actually convinced me that I could get a Ph.D. Finally, I would like to thank my family for their continual support and belief in my ability to succeed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure | Page

KEY TO SYMBOLS

Symbol/Definition	Term
$\boldsymbol{a} = (a_1, \ldots, a_s)'$	Vector
$\|\boldsymbol{a}\| = \sqrt{a_1^2 + \cdots + a_s^2}$	Euclidean norm of $\boldsymbol{a}$
$\boldsymbol{A} = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_t) = \{a_{ij}\}_{s \times t}$	Matrix
$\boldsymbol{A}' = \{a_{ji}\}_{t \times s}$	Transpose of $\boldsymbol{A}$
$\|\boldsymbol{A}\|$	Determinant of $\boldsymbol{A}$
$\mathrm{tr}(\boldsymbol{A})$	Trace of $\boldsymbol{A}$
$\mathrm{vec}\,(\boldsymbol{A}) = (\boldsymbol{a}_1', \ldots, \boldsymbol{a}_t')'$	Vector of $\boldsymbol{A}$
$\boldsymbol{A} \otimes \boldsymbol{B} = \begin{pmatrix} a_{11}\boldsymbol{B} & a_{12}\boldsymbol{B} \\ a_{21}\boldsymbol{B} & \ddots \end{pmatrix}$	Direct product
$\boldsymbol{A} \oplus \boldsymbol{B} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{B} \end{pmatrix}$	Direct sum
$\mathbb{R}$	Real numbers
Ω_p	Unit hypersphere of dimension p
∇	Gradient operator
$\sim$	Distributed as
iid	Independent and identically distributed
cdf	Cumulutave distribution function

Symbol/Definition	Term
$\mathrm{E}[\cdot]$	Expectation
$\mathrm{Cov}[\cdot,\cdot]$	Covariance
$\mathrm{V}[\cdot]$	Variance
$\xrightarrow{d}$	Converge in distribution
$\xrightarrow{p}$	Converge in probability
$AN(\mu,\sigma^2)$	Asymptotically normal r.v. with mean μ and variance σ^2
$\chi_k^2(\lambda)$	Chi-square r.v. with k d.f. and noncentrality parameter λ
ARE	Asymptotic relative efficiency
$o_p(f(n))$	Term when divided by $f(n)$ converges to zero in probability as $n \to \infty$
$O_p(f(n))$	Term when divided by $f(n)$ is bounded in probability as $n \to \infty$

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy

A NEW NONPARAMETRIC TEST FOR INDEPENDENCE BETWEEN TWO SETS OF VARIATES

By

Peter William Gieser

December 1993

Chairman: Ronald H. Randles
Major department: Statistics

A new nonparametric sign statistic based on interdirections is proposed for testing whether two sets of variates are independent. This interdirection quadrant statistic reduces to the sample coefficient of medial correlation (or quadrant statistic) when the two sets of variates have one variable each. It has an intuitive invariance property for the present problem and has a limiting chi-square distribution under the null hypothesis of independence when each set of variates is elliptically symmetric, making it asymptotically distribution-free. The new statistic is compared to the classical normal theory competitor, Wilks' likelihood ratio criterion, and a component-wise quadrant statistic. Using a novel model of dependence between the sets of variates enables the computation of Pitman asymptotic relative efficiencies (ARE's). The Pitman ARE's indicate that the interdirection quadrant statistic compares favorably to Wilks' likelihood ratio criterion when the sets of variates have heavy-tailed distributions and is uniformly better than the component-wise quadrant statistic when the sets of variates are spherically symmetric. A simulation study demonstrates the relative performances of the three competitors as well as some other statistics often

found in commercial software packages. The results indicate that the interdirection quadrant statistic performs better than the others for heavy-tailed distributions and is competitive for distributions with moderate tail weights. Finally, several applications of the interdirection quadrant statistic are illustrated, with comparisons to its competitors. In one example, the interdirection quadrant statistic's resistance to outliers is demonstrated.

CHAPTER 1
INTRODUCTION

1.1 Bivariate Tests

The question of whether the pair of random variables (X, Y) are stochastically independent, based on the random sample $\{(X_i, Y_i),\ i = 1, \ldots, n\}$ from a continuous distribution with density function $h(x, y)$, has generated vast amounts of research over the past century. After the "discovery" of the correlation concept by Galton (1888), many bivariate measures of correlation were invented to explore the nature of the dependency between X and Y. Examples include the classical Pearson product moment correlation coefficient (Pearson, 1896)

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\left(\sum_{i=1}^{n}(X_i - \bar{X})^2 \sum_{i=1}^{n}(Y_i - \bar{Y})^2\right)^{1/2}},$$

and numerous rank correlation statistics based on the ranks of the X_i's (Y_i's) denoted by $R_1, \ldots, R_n$ ($Q_1, \ldots, Q_n$) such as Spearman's rho (Spearman, 1904)

$$\rho = \frac{12}{n^3 - n} \sum_{i=1}^{n} \left(R_i - \frac{n+1}{2}\right)\left(Q_i - \frac{n+1}{2}\right),$$

Kendall's tau (Greiner, 1909; Kendall, 1938)

$$\tau = \frac{1}{n(n-1)} \sum_{i<j}^{n} \operatorname{sgn}(R_i - R_j) \operatorname{sgn}(Q_i - Q_j),$$

and the sample coefficient of medial correlation, or more simply the quadrant statistic (owing to the fact that it is based on the number of points in the four quadrants

defined by the marginal medians) (Blomqvist, 1950)

$$q' = \frac{1}{n}\sum_{i=1}^{n} \operatorname{sgn}\left(R_i - \frac{n+1}{2}\right) \operatorname{sgn}\left(Q_i - \frac{n+1}{2}\right)$$

$$= \frac{1}{n}\sum_{i=1}^{n} \operatorname{sgn}\left(X_i - \widetilde{X}\right) \operatorname{sgn}\left(Y_i - \widetilde{Y}\right),$$

where $\widetilde{X}$ ($\widetilde{Y}$) is the median of the X_i's (Y_i's). These measures can be used to conduct hypothesis tests for independence of X and Y. Many results are already known about the comparison of such (suitably normalized) bivariate tests of independence, and for various models of dependence, asymptotic relative efficiencies (ARE's) have been computed. We present a brief summary of these results.

Farlie (1960) introduces the bivariate distribution function

$$H(x,y) = F(x)G(y)\{1 + \alpha A(x)B(y)\}, \tag{1.1}$$

where $F(x)$ and $G(y)$ are the known marginal distribution functions of X and Y. For (1.1) to be a bonafide cdf, $d(F(x)A(x))/F(x)$, $d(G(y)B(y))/G(y)$, $A(x)$, and $B(y)$ need to be bounded, with $A(\infty) = B(\infty) = 0$. Clearly α is a measure of dependence, with $\alpha = 0$ denoting the independence of X and Y. The legal range of α is determined by the greatest and least values of the product $\{d(F(x)A(x))/F(x)\} \times \{d(G(y)B(y))/G(y)\}$ over all variation of x and y. Farlie (1961) then derives the asymptotic efficiency for a generalized correlation coefficient Γ devised by Daniels (1944). To each ordered pair of X's, (X_i, X_j), assign a score a_{ij} such that $a_{ij} = -a_{ji}$ and $a_{ii} = 0$. In a similar way assign scores b_{ij} using the Y's. The definition of Γ is then

$$\Gamma = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}b_{ij}}{\left(\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^2 \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}^2\right)^{1/2}}.$$

Gamma includes many well-known correlation coefficients as special cases, among which are ρ, τ, and q'. To get τ for example, put $a_{ij} = \operatorname{sgn}(R_i - R_j)$.

Farlie shows that for alternatives with the distributional form (1.1), there is always some coefficient of Daniels' family of coefficients that is fully efficient (i.e., the Pitman ARE with the maximum likelihood estimator is unity). We emphasize that the ARE's are determined by letting $\alpha \to 0$ as $n \to \infty$ and hence are Pitman-type ARE's. Here α is the parameter in (1.1) and not the level of the tests. Table 1.1 summarizes his results when $H(x, y)$ is the bivariate normal distribution. Note that in this case the maximum likelihood estimator is r, and α is the "true" correlation between X and Y. Farlie further notes that his results agree with those obtained by Blomqvist (1950) for q' and by Stuart (1954) for ρ and τ.

Table 1.1. Pitman ARE's Reported by Farlie Under Bivariate Normality

ARE(ρ, r)	$9/\pi^2$
ARE(τ, r)	$9/\pi^2$
ARE(q', r)	$4/\pi^2$

Konijn (1954) considers the model of dependence $X = \lambda_1 U + \lambda_2 V$ and $Y = \lambda_3 U + \lambda_4 V$, with U and V independent and λ_1, λ_2, λ_3, and λ_4 constants. The independence case is produced by setting $\lambda_2 = \lambda_3 = 0$. When U and V have the same marginal distribution G, he computes Pitman ARE's for several choices of G by letting $\lambda_2 \to 0$ and $\lambda_3 \to 0$ as $n \to \infty$. His results are summarized in Table 1.2.

Table 1.2. Pitman ARE's Computed by Konijn

	Normal	Uniform	Parabolic	Laplace
ARE(ρ, r)	$9/\pi^2$	1	0.8569	1.2656
ARE(τ, r)	$9/\pi^2$	1	0.8569	1.2656
ARE(q', r)	$4/\pi^2$	1/4	0.3164	1

Hájek and Šidák (1967) use the model $X = X^* + \Delta Z$ and $Y = Y^* + \Delta Z$ where X^*, Y^*, and Z are mutually independent with densities f, g, and m, respectively, and Δ is a positive constant. They require f and g be known and that $0 < \mathrm{Var}[Z] < \infty$. They focus on determining locally most powerful rank tests. When f and g are of logistic type, the locally most powerful rank test is based on ρ, and they indicate that q' is the resulting statistic when using approximate scores in the locally most powerful rank test for f and g of double exponential type. They acknowledge that they were unable to derive Pitman ARE's as $\Delta \to 0$, but Puri and Sen (1971) successfully complete the work using a model that encompasses Hájek and Šidák's as a special case. The general multivariate version of this model is discussed in Section 3.3. In the bivariate case, denote the joint density function of X and Y by $h_\Delta(x,y) = \int f(x - \Delta z)g(y - \Delta z)m(z)dz$. Let X, Y, and Z each have unit variances and define a sequence of alternatives by $\Delta_n = n^{-1/4}\Delta_0$. They show that the Pitman ARE of the statistic $T_n = n^{-1}\sum_{\alpha=1}^{n} J(F_n(X_\alpha))J(G_n(Y_\alpha))$ compared to r, where J is a standardized score function and F_n and G_n are the empirical cdf's of X and Y, respectively, is given by

$$\mathrm{ARE}(T_n, r) = \lim_{n\to\infty} \left[\Delta_n^{-1} \iint J(F(x))J(G(y))\, dH_{\Delta_n}(x,y)\right]^2 .$$

As an example, using the score function J_0 defined as

$$J_0(u) = \begin{cases} 1 & \text{if } 1/2 < u \leq 1, \\ 0 & \text{if } u = 1/2, \\ -1 & \text{if } 0 \leq u < 1/2 \end{cases}, \tag{1.2}$$

and assuming H_{Δ_n} is bivariate normal, we have that

$$\mathrm{ARE}(q', r) = \lim_{n\to\infty} \left[2/\pi \cdot \Delta_n^{-1} \sin^{-1}(\Delta_n/(1+\Delta_n))\right]^2 = 4/\pi^2.$$

In a series of exercises, Puri and Sen also investigate a slight variation of the Hájek and Šidák model in the bivariate case. They use $X = (1-\Delta)X^* + \Delta Z$ and $Y = (1-\Delta)Y^* + \Delta Z$ with X^*, Y^*, and Z mutually independent. Their modifications,

in conjunction with the model used by Konijn, were admittedly the inspiration for Model 1 considered in Section 3.2.

1.2 Multivariate Tests

Our interest, however, lies in the multivariate extension of this problem. Specifically, instead of testing whether two real-valued random variables are independent, we consider testing whether two vector-valued random variables are independent, where the dimensions of the vectors need not be the same. Let $\{\boldsymbol{X}_i \equiv (\boldsymbol{X}_i^{(1)\prime}, \boldsymbol{X}_i^{(2)\prime})', \ i = 1, \ldots, n\}$ be a random sample of n pairs of vectors from a continuous distribution with density function $f_X\left(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}\right)$, where $\boldsymbol{X}_i^{(k)}$ is $r_k \times 1$ and has marginal density $f_k\left(\boldsymbol{x}^{(k)}\right)$, $k = 1, 2$ (further occurrences of the index k generally mean the statement holds for both $k = 1$ and $k = 2$). Further, assume that $f_k\left(\boldsymbol{x}^{(k)}\right)$ represents a distribution that is elliptically symmetric and centered at the $r_k \times 1$ vector $\boldsymbol{\theta}_k$ (i.e., $f_k\left(\boldsymbol{x}^{(k)}\right)$ is a function of $ell = (\boldsymbol{x}^{(k)} - \boldsymbol{\theta}_k)'\Sigma_k^{-1}(\boldsymbol{x}^{(k)} - \boldsymbol{\theta}_k)$ alone, where Σ_k is a positive-definite matrix.) This is a common assumption in multivariate theory as it implies the simple structure that observations on ellipsoids defined by ell = constant are equally likely. The elliptically symmetric class is sufficiently general to accommodate a wide variety of different distributions, so that this assumption is not overly restrictive. Without loss of generality, we also take $r_1 \leq r_2$. We are interested in testing $H_0\colon f_X\left(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}\right) = f_1\left(\boldsymbol{x}^{(1)}\right) f_2\left(\boldsymbol{x}^{(2)}\right)$ versus $H_1 : \boldsymbol{X}_i^{(1)}$ and $\boldsymbol{X}_i^{(2)}$ are *correlated.*

1.2.1 Likelihood Ratio Criterion

Wilks (1935) derived the likelihood ratio criterion for testing $H_0\colon \Sigma_{12} = \mathbf{0}$ when the $\boldsymbol{X}_i$'s are multivariate normal with mean vector $\boldsymbol{\mu} = (\boldsymbol{\mu}_1', \boldsymbol{\mu}_2')'$ and covariance matrix $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}' & \Sigma_{22} \end{pmatrix}$. The normal distribution is unique, in that independence

is completely determined by the form of the covariance matrix. Thus $\Sigma_{12} = \mathbf{0}$ is equivalent to the null hypothesis of independence. If $\boldsymbol{A} = \sum_{\alpha=1}^{n}(\boldsymbol{X}_\alpha - \bar{\boldsymbol{X}})(\boldsymbol{X}_\alpha - \bar{\boldsymbol{X}})'$, and we partition it into $\boldsymbol{A}_{ij} = \sum_{\alpha=1}^{n}(\boldsymbol{X}_\alpha^{(i)} - \bar{\boldsymbol{X}}^{(i)})(\boldsymbol{X}_\alpha^{(j)} - \bar{\boldsymbol{X}}^{(j)})'$, $i, j = 1, 2$, then the criterion is expressed as

$$\begin{aligned} V^{n/2} &= \left[\frac{|\boldsymbol{A}|}{|\boldsymbol{A}_{11}||\boldsymbol{A}_{22}|}\right]^{n/2} \\ &= |\boldsymbol{I}_{r_1} - \boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{12}'|^{n/2} \\ &= |\boldsymbol{I}_{r_2} - \boldsymbol{A}_{12}'\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}|^{n/2}. \end{aligned}$$

The criterion also has a convenient limiting distribution under the null hypothesis, in that $-2\log V^{n/2} = -n\log V \xrightarrow{d} \chi^2_{r_1 r_2}$.

Muirhead (1982) shows that under the group of transformations described by

$$\mathcal{G} = \{g(\boldsymbol{B}, \boldsymbol{c}) \mid g(\boldsymbol{B}, \boldsymbol{c})(\boldsymbol{X}) = \boldsymbol{B}\boldsymbol{X} + \boldsymbol{c}\}$$

where $\boldsymbol{B} = \boldsymbol{B}_1 \oplus \boldsymbol{B}_2$, ($\boldsymbol{B}_k$ nonsingular $r_k \times r_k$) and $\boldsymbol{c} \in \mathbb{R}^{r_1+r_2}$, a maximal invariant is the set of sample canonical correlations $(\hat{\rho}_1, \ldots, \hat{\rho}_{r_1})$ where $\hat{\rho}_1^2 > \cdots > \hat{\rho}_{r_1}^2$ are the eigenvalues of $\boldsymbol{S} \equiv \boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{12}'$. The usefulness of the canonical correlations arises from the fact that through linear transformations on $\boldsymbol{X}^{(1)}$ and $\boldsymbol{X}^{(2)}$, the correlation structure implicit in the covariance matrix Σ can be reduced to a form involving only these parameters. In fact, the null hypothesis can be restated as $H_0\colon \rho_1 = \cdots = \rho_{r_1} = 0$. Note that the *nonzero* eigenvalues of $\boldsymbol{A}_{12}'\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}$ are identical to the eigenvalues of $\boldsymbol{S}$, so that the criterion is (as one would expect) invariant to the labeling of the two partitions of $\boldsymbol{X}$. Since we can rewrite V as $\prod_{i=1}^{r_1}(1 - \hat{\rho}_i^2)$, this shows that V is invariant under the group $\mathcal{G}$. When r_1 is one, the lone sample canonical correlation is just the multiple correlation coefficient between the $X_i^{(1)}$s and the $\boldsymbol{X}_i^{(2)}$s. Then if r_2 is also one, V is just $(1 - r^2)$, the sample coefficient of alienation. Thus, the likelihood ratio criterion is the sample vector coefficient of alienation and

a multivariate extension of the bivariate test of independence based on r. The invariance of V under the group $\mathcal{G}$ is an important property since it implies that a test using V will not depend on the underlying covariance structure of either the $\boldsymbol{X}_i^{(1)}$'s or $\boldsymbol{X}_i^{(2)}$'s. Other statistics that are also functions of the eigenvalues of $\boldsymbol{S}$ (squares of the sample canonical correlations) and hence invariant under $\mathcal{G}$, will be described in Chapter 4.

1.2.2 Component-wise Quadrant Statistic

A nonparametric approach to the problem is explored in Puri and Sen (1971), where a class of association parameters (and their sample counterparts) based on component-wise ranking is defined. The statistic they propose is computed using the elements of the matrix $\boldsymbol{T} = \begin{pmatrix} \boldsymbol{T}_{11} & \boldsymbol{T}_{12} \\ \boldsymbol{T}_{21} & \boldsymbol{T}_{22} \end{pmatrix}$. The elements of $\boldsymbol{T}_{ij} = \{T_{s_i s_j}^{(ij)}\}_{r_i \times r_j}$ are given by

$$T_{s_i s_j}^{(ij)} = \frac{1}{n} \sum_{\alpha=1}^{n} J\left(\frac{R_{s_i \alpha}^{(i)}}{n+1}\right) J\left(\frac{R_{s_j \alpha}^{(j)}}{n+1}\right).$$

Here, $R_{s_k \alpha}^{(k)}$ is the rank of $X_{s_k \alpha}^{(k)}$ among $X_{s_k 1}^{(k)}, \ldots, X_{s_k n}^{(k)}$ and J represents an arbitrary (standardized) score function. Puri and Sen base their test of independence on the statistic

$$S^J = \left[\frac{|\boldsymbol{T}|}{|\boldsymbol{T}_{11}||\boldsymbol{T}_{22}|}\right],$$

which is clearly analogous to V except the matrix $\boldsymbol{T}$ is used instead of the matrix $\boldsymbol{A}$. They also show that under the null hypothesis, $-n \log S^J \xrightarrow{d} \chi^2_{r_1 r_2}$, so that their procedure is a natural competitor of $-n \log V$.

A score function that will be of interest to us is J_0, defined in (1.2). With this score function, the elements of $\boldsymbol{T}_{12}$ are

$$T_{s_1 s_2}^{(12)} = \frac{1}{n} \sum_{\alpha=1}^{n} \operatorname{sgn}\left(X_{s_1 \alpha}^{(1)} - \widetilde{X}_{s_1}^{(1)}\right) \operatorname{sgn}\left(X_{s_2 \alpha}^{(2)} - \widetilde{X}_{s_2}^{(2)}\right),$$

so that S^{J_0} is a multivariate extension of q'. A problem with statistics based on the component-wise ranking scheme is that they do not have the desirable invariance property exhibited by V. Thus S^{J_0}, using component-wise ranking, fails to be invariant under the group $\mathcal{G}$, and its performance will be influenced by the (presumably) unknown underlying covariance structure of the $\boldsymbol{X}_i^{(1)}$'s and $\boldsymbol{X}_i^{(2)}$'s.

In Chapter 2, we propose a nonparametric competitor to V that like S^{J_0} is a multivariate extension of q', but maintains the important invariance discussed in Section 1.2. We derive the limiting null distribution in the case when $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ are known. Sufficient conditions for asserting that the limiting distribution is unchanged when $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ are replaced by the estimates $\hat{\boldsymbol{\theta}}_1$ and $\hat{\boldsymbol{\theta}}_2$ are given as well. In Chapter 3, we introduce a new model of dependence, with computations leading to the limiting distributions of the competing statistics under a contiguous sequence of alternatives based on this model. Pitman ARE's are calculated using these limiting results. In Chapter 4 we present a Monte Carlo study that corroborates empirically the results of Chapter 3, and in Chapter 5, we apply the new test and its competitors to some real-world data. We conclude with some general comments and indicate potential areas of future research in Chapter 6.

CHAPTER 2
INTERDIRECTION QUADRANT STATISTIC

2.1 Definition

Let $\hat{\boldsymbol{\theta}}_1$ based on $\boldsymbol{X}_1^{(1)}, \ldots, \boldsymbol{X}_n^{(1)}$ and $\hat{\boldsymbol{\theta}}_2$ based on $\boldsymbol{X}_1^{(2)}, \ldots, \boldsymbol{X}_n^{(2)}$ be equivariant (under the group $\mathcal{G}$) estimators of $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$, respectively, such that both $(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1)$ and $(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_2)$ are $O_p(n^{-1/2})$. The interdirection quadrant statistic is defined as

$$\hat{Q}_n(\hat{\boldsymbol{\theta}}_1, \hat{\boldsymbol{\theta}}_2) = \frac{r_1 r_2}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \cos(\pi \hat{p}_1(\boldsymbol{X}_i^{(1)}, \boldsymbol{X}_j^{(1)}; \hat{\boldsymbol{\theta}}_1)) \cos(\pi \hat{p}_2(\boldsymbol{X}_i^{(2)}, \boldsymbol{X}_j^{(2)}; \hat{\boldsymbol{\theta}}_2)),$$

where $\hat{p}_k(\boldsymbol{X}_i^{(k)}, \boldsymbol{X}_j^{(k)}; \boldsymbol{\gamma}_k)$ is the interdirection proportion, first defined by Randles (1989), between $(\boldsymbol{X}_i^{(k)} - \boldsymbol{\gamma}_k)$ and $(\boldsymbol{X}_j^{(k)} - \boldsymbol{\gamma}_k)$. We now describe how to calculate this interdirection proportion. If we let $\boldsymbol{Z}_i^{(k)} = \boldsymbol{X}_i^{(k)} - \boldsymbol{\gamma}_k$, then the number of hyperplanes defined by the origin and $r_k - 1$ other $\boldsymbol{Z}^{(k)}$s (not $\boldsymbol{Z}_i^{(k)}$ or $\boldsymbol{Z}_j^{(k)}$) such that $\boldsymbol{Z}_i^{(k)}$ and $\boldsymbol{Z}_j^{(k)}$ are on opposite sides, is called the interdirection count. To get the interdirection proportion, divide by the total number of hyperplanes considered. Note that this is the simpler and more natural divisor given by Randles. He makes a small sample adjustment in this divisor so that his interdirection sign test is equivalent to Blumen's test in the bivariate setting. The interdirection count measures the angular distance between $\boldsymbol{Z}_i^{(k)}$ and $\boldsymbol{Z}_j^{(k)}$ relative to the origin and the position of the other $\boldsymbol{Z}^{(k)}$s. Randles showed that under a wide class of population models (distributions with elliptical directions, a superset of elliptically symmetric distributions), when the $\boldsymbol{X}^{(k)}$s are centered on the symmetry point of the distribution, $\boldsymbol{\theta}_k$, the interdirection count (i) is invariant under nonsingular linear transformations, (ii) uses only the direction of

each $(\boldsymbol{X}^{(k)} - \boldsymbol{\theta}_k)$ from the origin, and (iii) has a distribution-free property. Properties (i) and (ii) continue to hold when centering about $\hat{\boldsymbol{\theta}}_k$, because of its equivariance under $\mathcal{G}$ but property (iii) fails to hold. Then, since the interdirection proportions are invariant under the group $\mathcal{G}$, $\hat{Q}_n(\hat{\boldsymbol{\theta}}_1, \hat{\boldsymbol{\theta}}_2)$ is as well. Notice that if $r_1 = r_2 = 1$ (the bivariate case), then

$$\cos(\pi \hat{p}_k(X_i^{(k)}, X_j^{(k)}; \hat{\theta}_k)) = \operatorname{sgn}\left(X_i^{(k)} - \hat{\theta}_k\right) \operatorname{sgn}\left(X_j^{(k)} - \hat{\theta}_k\right),$$

so that

$$\hat{Q}_n(\hat{\theta}_1, \hat{\theta}_2) = \frac{1}{n}\left[\sum_{i=1}^{n} \operatorname{sgn}\left(X_i^{(1)} - \hat{\theta}_1\right) \operatorname{sgn}\left(X_i^{(2)} - \hat{\theta}_2\right)\right]^2,$$

which implies that $\hat{Q}_n(\widetilde{X}^{(1)}, \widetilde{X}^{(2)}) = (\sqrt{n}q')^2$. Thus, as the name indicates, $\hat{Q}_n(\hat{\theta}_1, \hat{\theta}_2)$ is an extension of the bivariate quadrant statistic. In the next section, we derive the limiting null distribution for the interdirection quadrant statistic in the case where $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ are assumed known.

2.2 Null Distribution When $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ Is Known

In establishing the limiting null distribution of $\hat{Q}_n(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ we first seek the limiting null distribution of a simpler approximating quantity. The simplicity derives from the fact that the interdirection function $\hat{p}_k(\cdot, \cdot; \boldsymbol{\theta}_k)$ is replaced by its expected value $p_k(\cdot, \cdot; \boldsymbol{\theta}_k)$ under the null hypothesis. All calculations in this section are assumed to be done under the null hypothesis, so no further explicit mention of this fact will be made. Because $\hat{Q}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ is invariant under the group $\mathcal{G}$, without loss of generality we assume that $f_k\left(\boldsymbol{x}^{(k)}\right)$ represents a distribution that is spherically symmetric and centered at the origin (i.e., $\boldsymbol{\theta}_k = \boldsymbol{0}$ and $\Sigma_k = \boldsymbol{I}_{r_k}$.) Define

$$Q_n = \frac{r_1 r_2}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \cos(\pi p_1(i,j)) \cos(\pi p_2(i,j)),$$

where we have kept just the subscripts i and j, since the subscript k on $p_k(i,j)$ (and also $\hat{p}_k(i,j)$) indicates whether it is a function of the $\boldsymbol{X}^{(1)}$'s or $\boldsymbol{X}^{(2)}$'s, and

$$\begin{aligned}
p_k(i,j) &= \mathrm{E}_{H_0}\left[\hat{p}_k(i,j) \mid \boldsymbol{X}_i^{(k)}, \boldsymbol{X}_j^{(k)}\right] \\
&= \mathrm{P}\left[\begin{array}{c} \text{a hyperplane defined by } r_k - 1 \\ \boldsymbol{X}^{(k)}\text{'s and the origin lies between} \\ \boldsymbol{X}_i^{(k)} \text{ and } \boldsymbol{X}_j^{(k)} \end{array} \middle| \, \boldsymbol{X}_i^{(k)}, \boldsymbol{X}_j^{(k)}\right] \\
&= (\text{Angle between } \boldsymbol{X}_i^{(k)} \text{ and } \boldsymbol{X}_j^{(k)})/\pi.
\end{aligned}$$

The last equality follows because of the underlying spherical symmetry. Let $\boldsymbol{X}^{(k)} = R^{(k)}\boldsymbol{U}^{(k)}$, where $R^{(k)} = \|\boldsymbol{X}^{(k)}\|$ and $\boldsymbol{U}^{(k)} = \boldsymbol{X}^{(k)}/\|\boldsymbol{X}^{(k)}\|$. Spherical symmetry guarantees that $R_1^{(k)}, \ldots, R_n^{(k)}$ are positive quantities independent of $\boldsymbol{U}_1^{(k)}, \ldots, \boldsymbol{U}_n^{(k)}$, and the $\boldsymbol{U}_i^{(k)}$s are iid uniform on the unit hypersphere of dimension r_k. Here $\boldsymbol{U}^{(k)}$ is the direction of $\boldsymbol{X}^{(k)}$, and it is easy to see that both $\hat{p}_k(i,j)$ and $p_k(i,j)$ depend only on the directions $\boldsymbol{U}_i^{(k)}$ and $\boldsymbol{U}_j^{(k)}$. Now

$$\cos(\pi p_k(i,j)) = \cos(\text{angle between } \boldsymbol{U}_i^{(k)} \text{ and } \boldsymbol{U}_j^{(k)}) = \boldsymbol{U}_i^{(k)\prime}\boldsymbol{U}_j^{(k)},$$

so that

$$\begin{aligned}
Q_n &= \frac{r_1 r_2}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} \boldsymbol{U}_i^{(1)\prime}\boldsymbol{U}_j^{(1)}\boldsymbol{U}_i^{(2)\prime}\boldsymbol{U}_j^{(2)} \\
&= \frac{1}{n}\sum_{s=1}^{r_1}\sum_{t=1}^{r_2}\left[\sum_{\alpha=1}^{n}\sqrt{r_1 r_2}\,U_{s\alpha}^{(1)}U_{t\alpha}^{(2)}\right]^2.
\end{aligned} \tag{2.1}$$

Some moments involving $\boldsymbol{U}^{(1)}$ and $\boldsymbol{U}^{(2)}$ will be useful in subsequent derivations. If $\boldsymbol{U}^{(1)}$ and $\boldsymbol{U}^{(2)}$ are independent with $\boldsymbol{U}^{(1)} \sim \text{Uniform}(\Omega_{r_1})$ and $\boldsymbol{U}^{(2)} \sim \text{Uniform}(\Omega_{r_2})$,

then it is easy to show that

$$
\begin{aligned}
\mathrm{E}_{H_0}\left[\boldsymbol{U}^{(k)}\right] &= \mathbf{0} \\
\mathrm{E}_{H_0}\left[\boldsymbol{U}^{(k)}\boldsymbol{U}^{(k)\prime}\right] &= \frac{1}{r_k}\boldsymbol{I}_{r_k} \\
\mathrm{E}_{H_0}\left[\boldsymbol{U}^{(k)\prime}\boldsymbol{M}\boldsymbol{U}^{(k)}\right] &= \frac{1}{r_k}\mathrm{vec}\,(\boldsymbol{M})'\,\mathrm{vec}\,(\boldsymbol{I}_{r_k}) \\
\mathrm{E}_{H_0}\left[\boldsymbol{U}^{(1)\prime}\boldsymbol{M}_1\boldsymbol{U}^{(2)}\boldsymbol{U}^{(1)\prime}\boldsymbol{M}_2\boldsymbol{U}^{(2)}\right] &= \frac{1}{r_1 r_2}\mathrm{vec}\,(\boldsymbol{M}_1)'\,\mathrm{vec}\,(\boldsymbol{M}_2)\,.
\end{aligned}
\tag{2.2}
$$

Two basic classical limit results are restated here for future reference.

<u>Lemma 2.2.1</u> $\boldsymbol{a}'\boldsymbol{Z}_n \sim AN(\boldsymbol{a}'\boldsymbol{\mu}, \boldsymbol{a}'\Sigma\boldsymbol{a})$ for any $\boldsymbol{a} \neq \mathbf{0}$ if and only if $\boldsymbol{Z}_n \sim AN(\boldsymbol{\mu}, \Sigma)$.

<u>Proof of Lemma 2.2.1</u> (See Serfling, 1980, p. 18). □

<u>Lemma 2.2.2</u> If $\boldsymbol{Z}_n \sim AN(\boldsymbol{\mu}, \boldsymbol{I}_k)$, then $\boldsymbol{Z}_n'\boldsymbol{Z}_n \xrightarrow{d} \chi_k^2(\boldsymbol{\mu}'\boldsymbol{\mu})$.

<u>Proof of Lemma 2.2.2</u> (See Serfling, 1980, p. 128). □

We now have all the necessary tools to derive the asymptotic distribution of the interdirection quadrant statistic. We begin by finding the limiting distribution of the approximating quantity Q_n.

<u>Theorem 2.2.1</u> $Q_n \xrightarrow{d} \chi^2_{r_1 r_2}$.

<u>Proof of Theorem 2.2.1</u> Let $\boldsymbol{B} = \{b_{st}\}_{r_1\times r_2}$ be an arbitrary matrix of constants that are not all zero and define $\boldsymbol{Z} = \{\sum_{\alpha=1}^n \sqrt{r_1 r_2}U_{s\alpha}^{(1)}U_{t\alpha}^{(2)}\}_{r_1\times r_2}$. Then

$$
\begin{aligned}
\mathrm{vec}\,(\boldsymbol{B})'\,\mathrm{vec}\,(\boldsymbol{Z}) &= \sum_{s=1}^{r_1}\sum_{t=1}^{r_2} b_{st}Z_{st} \\
&= \sum_{\alpha=1}^{n}\left[\sum_{s=1}^{r_1}\sum_{t=1}^{r_2}\sqrt{r_1 r_2}\,b_{st}U_{s\alpha}^{(1)}U_{t\alpha}^{(2)}\right] \\
&= \sum_{\alpha=1}^{n}\boldsymbol{U}_\alpha^{(1)\prime}(\sqrt{r_1 r_2}\boldsymbol{B})\boldsymbol{U}_\alpha^{(2)},
\end{aligned}
$$

which, using (2.2), is seen to be a sum of iid random variables with mean zero and variance $\mathrm{vec}(\boldsymbol{B})'\,\mathrm{vec}(\boldsymbol{B})$. Now $n^{-1/2}\mathrm{vec}(\boldsymbol{B})'\,\mathrm{vec}(\boldsymbol{Z}) \sim AN(0, \mathrm{vec}(\boldsymbol{B})'\,\mathrm{vec}(\boldsymbol{B}))$ via the central limit theorem. Lemmas 2.2.1 and 2.2.2 then imply that $n^{-1}\mathrm{vec}(\boldsymbol{Z})'\,\mathrm{vec}(\boldsymbol{Z}) \xrightarrow{d} \chi^2_{r_1 r_2}$. The result follows by noting that $n^{-1}\mathrm{vec}(\boldsymbol{Z})'\,\mathrm{vec}(\boldsymbol{Z}) = n^{-1}\sum_{s=1}^{r_1}\sum_{t=1}^{r_2} Z_{st}^2$, which, referring to (2.1), is just Q_n. □

Now using arguments similar to those in Randles (1989), we are able to find the limiting distribution of $\hat{Q}_n$.

<u>Theorem 2.2.2</u> $\hat{Q}_n \xrightarrow{d} \chi^2_{r_1 r_2}$.

<u>Proof of Theorem 2.2.2</u>

$$\mathrm{E}_{H_0}\left[(\hat{Q}_n - Q_n)^2\right]$$

$$= \mathrm{E}_{H_0}\left[\left(\frac{r_1 r_2}{n}\sum_{i,j}^{n}\cos(\pi\hat{p}_1(i,j))\cos(\pi\hat{p}_2(i,j)) - \cos(\pi p_1(i,j))\cos(\pi p_2(i,j))\right)^2\right]$$

$$= \frac{r_1^2 r_2^2}{n^2}\sum_{i,j,i',j'}^{n}\mathrm{E}_{H_0}\Big[\Big\{\cos(\pi\hat{p}_1(i,j))\cos(\pi\hat{p}_2(i,j)) - \cos(\pi p_1(i,j))\cos(\pi p_2(i,j))\Big\}$$

$$\times\Big\{\cos(\pi\hat{p}_1(i',j'))\cos(\pi\hat{p}_2(i',j')) - \cos(\pi p_1(i',j'))\cos(\pi p_2(i',j'))\Big\}\Big]$$

$$= \frac{2r_1^2 r_2^2 n(n-1)}{n^2}\mathrm{E}_{H_0}\Big[\Big\{\cos(\pi\hat{p}_1(i,j))\cos(\pi\hat{p}_2(i,j)) - \cos(\pi p_1(i,j))\cos(\pi p_2(i,j))\Big\}^2\Big],$$

where the last equality is seen by considering the following. Let $\boldsymbol{U}_i^{(1)} = D_i\boldsymbol{A}_i$, where $D_i = \mathrm{sgn}\left(U_{i1}^{(1)}\right)$ and $\boldsymbol{A}_i = \mathrm{sgn}\left(U_{i1}^{(1)}\right)\boldsymbol{U}_i^{(1)}$. Then $\boldsymbol{A}_i$ shows the observed axis and D_i indicates which end of the axis was observed. Note that D_i is independent of $\boldsymbol{A}_i$ and $D_1, \ldots, D_n$ are iid Bernoulli random variables with probability of success 1/2. Taking the expectation first with respect to the D's, when one or more of the four subscripts is unique, the expected value is zero. Further, if $i = j$ or $i' = j'$ the integrand is zero since $\hat{p}_k(i,i) = p_k(i,i) = 0$. Now the last expectation converges to zero because

Randles has shown that $\hat{p}_k(i,j) = p_k(i,j) + o_p(1)$, and the integrand is bounded so the Lebesgue Dominated Convergence Theorem can be applied. Thus, $\hat{Q}_n = Q_n + o_p(1)$, and Theorem 2.2.1 yields the desired result. □

We have now established that $\hat{Q}_n$ has a convenient asymptotic null distribution, which makes it a viable option for the present hypothesis testing scenario. Further, since it has the identical limiting null distribution as $-n \log V$ and $-n \log S^{J_0}$, the relative performance of the three competitors can be fairly measured. However, in practice the values of $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ are rarely known so that they must be estimated. The next section considers this situation.

2.3 Null Distribution When $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ Is Unknown

Unfortunately, when $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ is replaced by $(\hat{\boldsymbol{\theta}}_1, \hat{\boldsymbol{\theta}}_2)$, the proof of convergence to the $\chi^2_{r_1 r_2}$ distribution is much more difficult. We wish to find sufficient conditions under which $\hat{Q}_n(\hat{\boldsymbol{\theta}}_1, \hat{\boldsymbol{\theta}}_2) = \hat{Q}_n(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) + o_p(1)$. Since

$$
\begin{aligned}
&\hat{Q}_n(\hat{\boldsymbol{\theta}}_1, \hat{\boldsymbol{\theta}}_2) - \hat{Q}_n(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) \\
&= \frac{r_1 r_2}{n} \sum_{i,j}^{n} \left\{ \cos(\pi \hat{p}_1(i,j;\hat{\boldsymbol{\theta}}_1)) \cos(\pi \hat{p}_2(i,j;\hat{\boldsymbol{\theta}}_2)) - \cos(\pi \hat{p}_1(i,j;\boldsymbol{\theta}_1)) \cos(\pi \hat{p}_2(i,j;\boldsymbol{\theta}_2)) \right\} \\
&= \frac{r_1 r_2}{n} \sum_{i,j}^{n} \left\{ \cos(\pi \hat{p}_1(i,j;\hat{\boldsymbol{\theta}}_1)) \cos(\pi \hat{p}_2(i,j;\hat{\boldsymbol{\theta}}_2)) - \cos(\pi \hat{p}_1(i,j;\hat{\boldsymbol{\theta}}_1)) \cos(\pi \hat{p}_2(i,j;\boldsymbol{\theta}_2)) \right\} \\
&\quad + \frac{r_1 r_2}{n} \sum_{i,j}^{n} \left\{ \cos(\pi \hat{p}_1(i,j;\hat{\boldsymbol{\theta}}_1)) \cos(\pi \hat{p}_2(i,j;\boldsymbol{\theta}_2)) - \cos(\pi \hat{p}_1(i,j;\boldsymbol{\theta}_1)) \cos(\pi \hat{p}_2(i,j;\boldsymbol{\theta}_2)) \right\} \\
&= B_{1n} + B_{2n},
\end{aligned}
\tag{2.3}
$$

it suffices to show that both $B_{1n} \xrightarrow{p} 0$ and $B_{2n} \xrightarrow{p} 0$. Consider B_{1n} first. The strategy for dealing with B_{1n} is to show that it suffices for the second conditional moment of B_{1n} to converge in probability to zero. The conditional moment is useful in that

it "separates" B_{1n} into a sum of terms whose factors involve only the $\boldsymbol{X}^{(1)}$'s or only the $\boldsymbol{X}^{(2)}$'s. The limiting behavior of these factors is then established in Appendix A. More formally, let $\mathcal{X}_1 = \{\boldsymbol{X}_i^{(1)},\ i = 1, \ldots, n\}$, $\mathcal{X}_2 = \{\boldsymbol{X}_i^{(2)},\ i = 1, \ldots, n\}$, and $\mathcal{R} = \{R_i,\ i = 1, \ldots, n\}$, where $\mathcal{R}$ is a permutation of the integers $\{1, \ldots, n\}$. Conditional on $\mathcal{X}_1$ and $\mathcal{X}_2$, the only random component of B_{1n} is the way in which the $\boldsymbol{X}^{(1)}$'s and $\boldsymbol{X}^{(2)}$'s are matched. In other words, given $\mathcal{X}_1$ and $\mathcal{X}_2$, under the null hypothesis B_{1n} is a function of $\{(\boldsymbol{X}_{R_i}^{(1)}, \boldsymbol{X}_i^{(2)}),\ i = 1, \ldots, n\}$, where $\mathcal{R}$ is uniformly distributed over all permutations of $\{1, \ldots, n\}$. Given this setup, we prove the following lemma.

<u>Lemma 2.3.1</u> If $\mathrm{E}_{H_0}\left[B_{1n}^2 \mid \mathcal{X}_1, \mathcal{X}_2\right] \xrightarrow{p} 0$, then $B_{1n} \xrightarrow{p} 0$.

<u>Proof of Lemma 2.3.1</u> Since

$$B_{1n} = \frac{r_1 r_2}{n} \sum_{i,j}^{n} \cos(\pi \hat{p}_1(i,j;\hat{\boldsymbol{\theta}}_1)) \left\{\cos(\pi \hat{p}_2(i,j;\hat{\boldsymbol{\theta}}_2)) - \cos(\pi \hat{p}_2(i,j;\boldsymbol{\theta}_2))\right\},$$

$$G_n(\mathcal{X}_1, \mathcal{X}_2) \equiv \mathrm{E}_{H_0}\left[B_{1n}^2 \mid \mathcal{X}_1, \mathcal{X}_2\right] = \mathrm{E}_{H_0}\left[\left(\frac{r_1 r_2}{n} \sum_{i,j}^{n} c(R_i, R_j) d(i,j)\right)^2\right],$$

where

$$c(i,j) = \cos(\pi \hat{p}_1(i,j;\hat{\boldsymbol{\theta}}_1))$$

and

$$d(i,j) = \cos(\pi \hat{p}_2(i,j;\hat{\boldsymbol{\theta}}_2)) - \cos(\pi \hat{p}_2(i,j;\boldsymbol{\theta}_2)).$$

For $\epsilon > 0$ and $\delta > 0$,

$$\begin{aligned}
\mathrm{P}\left[|B_{1n}| > \epsilon\right] &= \mathrm{P}\left[|B_{1n}| > \epsilon,\ G_n(\mathcal{X}_1, \mathcal{X}_2) > \frac{\epsilon^2 \delta}{2}\right] + \mathrm{P}\left[|B_{1n}| > \epsilon,\ G_n(\mathcal{X}_1, \mathcal{X}_2) < \frac{\epsilon^2 \delta}{2}\right] \\
&\leq \mathrm{P}\left[G_n(\mathcal{X}_1, \mathcal{X}_2) > \frac{\epsilon^2 \delta}{2}\right] + \mathrm{E}_{H_0}\left[I\left(|B_{1n}| > \epsilon,\ G_n(\mathcal{X}_1, \mathcal{X}_2) < \frac{\epsilon^2 \delta}{2}\right)\right]
\end{aligned}$$

which for n sufficiently large,

$$
\begin{aligned}
&\leq \frac{\delta}{2} + \mathrm{E}_{H_0}\left[\mathrm{E}_{H_0}\left[I\left(|B_{1n}| > \epsilon,\ G_n(\mathcal{X}_1, \mathcal{X}_2) < \frac{\epsilon^2\delta}{2}\right) \middle| \mathcal{X}_1, \mathcal{X}_2\right]\right] \\
&= \frac{\delta}{2} + \mathrm{E}_{H_0}\left[I\left(G_n(\mathcal{X}_1, \mathcal{X}_2) < \frac{\epsilon^2\delta}{2}\right) \mathrm{P}\left[|B_{1n}| > \epsilon \mid \mathcal{X}_1, \mathcal{X}_2\right]\right] \\
&\leq \frac{\delta}{2} + \mathrm{E}_{H_0}\left[I\left(G_n(\mathcal{X}_1, \mathcal{X}_2) < \frac{\epsilon^2\delta}{2}\right) \frac{1}{\epsilon^2}\mathrm{E}_{H_0}\left[B_{1n}^2 \mid \mathcal{X}_1, \mathcal{X}_2\right]\right] \\
&= \frac{\delta}{2} + \mathrm{E}_{H_0}\left[I\left(G_n(\mathcal{X}_1, \mathcal{X}_2) < \frac{\epsilon^2\delta}{2}\right) \frac{1}{\epsilon^2}G_n(\mathcal{X}_1, \mathcal{X}_2)\right] \\
&\leq \frac{\delta}{2} + \mathrm{E}_{H_0}\left[\frac{1}{\epsilon^2}\frac{\epsilon^2\delta}{2}\right] \\
&= \delta. \qquad \square
\end{aligned}
$$

In light of Lemma 2.3.1, it suffices to show $G_n(\mathcal{X}_1, \mathcal{X}_2) \xrightarrow{\mathcal{P}} 0$. With this in mind, define

$$S = \frac{r_1 r_2}{n}\sum_{i,j}^{n} c(R_i, R_j)d(i,j) = \frac{r_1 r_2}{n}\sum_{i \neq j}^{n} c(R_i, R_j)d(i,j)$$

where the last equality follows since $d(i,i) = 0$. This means that $G_n(\mathcal{X}_1, \mathcal{X}_2) = \mathrm{E}_{H_0}\left[S^2\right]$, so that in order to write out the expression for $G_n(\mathcal{X}_1, \mathcal{X}_2)$ we need only find the second moment of S. First, we will need some preliminary results regarding the moments of $c(R_i, R_j)$. Let P_n represent the collection of all permutations of $\{1, \ldots, n\}$, ${}_nC_k$ the collection of all subsets of size k from $\{1, \ldots, n\}$, and ${}_nD_k$ the collection of all subsets of size k and their permutations from $\{1, \ldots, n\}$.

Clearly, for $i \neq j$,

$$\begin{aligned}
\mathrm{E}_{H_0}\left[c(R_i, R_j)\right] &= \sum_{\boldsymbol{\alpha} \in P_n} c(\alpha_i, \alpha_j) \mathrm{P}\left[\mathcal{R} = \boldsymbol{\alpha}\right] \\
&= \sum_{\boldsymbol{\beta} \in {}_nD_2} c(\beta_1, \beta_2) \mathrm{P}\left[R_i = \beta_1, R_j = \beta_2\right] \\
&= \frac{1}{n(n-1)} \sum_{\boldsymbol{\beta} \in {}_nD_2} c(\beta_1, \beta_2) \\
&\equiv \bar{c},
\end{aligned}$$

and in a completely analogous way we have for $i \neq j$ and $i' \neq j'$ that if $i' = i$ and $j' = j$, then

$$\begin{aligned}
\mathrm{Cov}_{H_0}\left[c(R_i, R_j), c(R_{i'}, R_{j'})\right] &= \mathrm{V}_{H_0}\left[c(i, j)\right] \\
&= \frac{1}{n(n-1)} \sum_{\boldsymbol{\beta} \in {}_nD_2} \left\{c(\beta_1, \beta_2) - \bar{c}\right\}^2 \\
&\equiv \frac{c_1}{n(n-1)},
\end{aligned}$$

and if $i' = i$ or $j' = j$ (but not both), then

$$\begin{aligned}
\mathrm{Cov}_{H_0}\left[c(R_i, R_j) c(R_{i'}, R_{j'})\right] &= \mathrm{Cov}_{H_0}\left[c(R_i, R_j) c(R_i, R_{j'})\right] \\
&= \frac{1}{n(n-1)(n-2)} \sum_{\boldsymbol{\beta} \in {}_nD_3} \left\{c(\beta_1, \beta_2) - \bar{c}\right\}\left\{c(\beta_1, \beta_3) - \bar{c}\right\} \\
&\equiv \frac{c_2}{n(n-1)(n-2)},
\end{aligned}$$

and if neither $i' = i$ nor $j' = j$, then

$$\begin{aligned}
&\mathrm{Cov}_{H_0}\left[c(R_i, R_j) c(R_{i'}, R_{j'})\right] \\
&= \frac{1}{n(n-1)(n-2)(n-3)} \sum_{\boldsymbol{\beta} \in {}_nD_4} \left\{c(\beta_1, \beta_2) - \bar{c}\right\}\left\{c(\beta_3, \beta_4) - \bar{c}\right\} \\
&\equiv \frac{c_3}{n(n-1)(n-2)(n-3)}.
\end{aligned}$$

Let $\bar{d}$, d_1, d_2, and d_3 be the analogous quantities in $d(i,j)$. Now using these expressions for $\mathrm{Cov}_{H_0}\left[c(R_i, R_j)c(R_i', R_j')\right]$, we see that

$$\mathrm{V}_{H_0}[S] = \mathrm{V}_{H_0}\left[\sum_{i\neq j}^{n} d(i,j)c(R_i, R_j)\right]$$

$$= \sum_{i\neq j}^{n}\sum_{i'\neq j'}^{n} d(i,j)d(i',j')\mathrm{Cov}_{H_0}\left[c(R_i, R_j)c(R_i', R_j')\right]$$

$$= \frac{c_1}{n(n-1)}\sum_{\boldsymbol{\beta}\in\, _nD_2}^{n} d^2(\beta_1,\beta_2) + \frac{4c_2}{n(n-1)(n-2)}\sum_{\boldsymbol{\beta}\in\, _nD_3} d(\beta_1,\beta_2)d(\beta_1,\beta_3)$$

$$+ \frac{c_3}{n(n-1)(n-2)(n-3)}\sum_{\boldsymbol{\beta}\in\, _nD_4} d(\beta_1,\beta_2)d(\beta_3,\beta_4),$$

and since

$$\mathrm{V}_{H_0}\left[\sum_{i\neq j}^{n} d(i,j)c(R_i, R_j)\right] = \mathrm{V}_{H_0}\left[\sum_{i\neq j}^{n}\{d(i,j)-\bar{d}\}c(R_i, R_j)\right],$$

we have that

$$\mathrm{V}_{H_0}[S] = \frac{c_1 d_1}{n(n-1)} + \frac{4c_2 d_2}{n(n-1)(n-2)} + \frac{c_3 d_3}{n(n-1)(n-2)(n-3)}.$$

Note also that since $\sum_{i\neq j}^{n}\{c(i,j)-\bar{c}\} = 0$,

$$\left(\sum_{i\neq j}^{n}\left\{c(i,j)-\bar{c}\right\}\right)^2 = c_1 + 4c_2 + c_3 = 0,$$

so that $c_3 = -(c_1 + 4c_2)$, where a similar result holds for the d's. Finally,

$$\mathrm{E}_{H_0}[S] = \mathrm{E}_{H_0}\left[\sum_{i\neq j}^{n} d(i,j)c(R_i, R_j)\right]$$

$$= \sum_{i\neq j}^{n} d(i,j)\mathrm{E}_{H_0}\left[c(R_i, R_j)\right]$$

$$= \bar{c}\sum_{i\neq j}^{n} d(i,j)$$

$$= n(n-1)\bar{c}\bar{d}.$$

Since $\mathrm{E}_{H_0}[S^2] = \mathrm{V}_{H_0}[S] + \mathrm{E}^2_{H_0}[S]$, $G_n(\mathcal{X}_1, \mathcal{X}_2)$ can be expressed as

$$G_n(\mathcal{X}_1, \mathcal{X}_2) = \frac{r_1^2 r_2^2}{n^2}\left[\frac{c_1 d_1}{n(n-1)} + \frac{4c_2 d_2}{n(n-1)(n-2)}\right.$$

$$\left. - \frac{(c_1 + 4c_2)(d_1 + 4d_2)}{n(n-1)(n-2)(n-3)} + \left(n(n-1)\bar{c}\bar{d}\right)^2\right].$$

To show $G_n(\mathcal{X}_1, \mathcal{X}_2) \xrightarrow{p} 0$ (and hence that $B_{1n} \xrightarrow{p} 0$), it suffices to show $\bar{c} = o_p(n^{-1/2})$, $c_1 = O(n^2)$, $c_2 = o_p(n^{5/2})$, $\bar{d} = o_p(n^{-1/2})$, $d_1 = o_p(n^2)$, and $d_2 = o_p(n^{5/2})$. As stated, these results, along with the requisite assumptions needed, are in Appendix A.

Recalling that $\hat{Q}_n(\hat{\boldsymbol{\theta}}_1, \hat{\boldsymbol{\theta}}_2) - \hat{Q}_n(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = B_{1n} + B_{2n}$ (see (2.3)), we must now show $B_{2n} \xrightarrow{p} 0$ to complete the argument that $\hat{Q}_n(\hat{\boldsymbol{\theta}}_1, \hat{\boldsymbol{\theta}}_2) = \hat{Q}_n(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) + o_p(1)$. Of course it suffices to show $\mathrm{E}_{H_0}[B_{2n}^2] \to 0$. Since

$$B_{2n} = \frac{r_1 r_2}{n} \sum_{i,j}^{n} \cos(\pi \hat{p}_2(i,j;\boldsymbol{\theta}_2)) \left\{\cos(\pi \hat{p}_1(i,j;\hat{\boldsymbol{\theta}}_1)) - \cos(\pi \hat{p}_1(i,j;\boldsymbol{\theta}_1))\right\},$$

we have

$$\mathrm{E}_{H_0}\left[B_{2n}^2\right] = \frac{r_1^2 r_2^2}{n^2} \sum_{i,j,i',j'}^{n} \mathrm{E}_{H_0}\left[\cos(\pi \hat{p}_2(i,j;\boldsymbol{\theta}_2)) \cos(\pi \hat{p}_2(i',j';\boldsymbol{\theta}_2))\right]$$

$$\times \mathrm{E}_{H_0}\left[\left\{\cos(\pi \hat{p}_1(i,j;\hat{\boldsymbol{\theta}}_1)) - \cos(\pi \hat{p}_1(i,j;\boldsymbol{\theta}_1))\right\}\right.$$

$$\left. \times \left\{\cos(\pi \hat{p}_1(i',j';\hat{\boldsymbol{\theta}}_1)) - \cos(\pi \hat{p}_1(i',j';\boldsymbol{\theta}_1))\right\}\right]$$

$$= \frac{2 r_1^2 r_2^2 n(n-1)}{n^2} \mathrm{E}_{H_0}\left[\cos^2(\pi \hat{p}_2(i,j;\boldsymbol{\theta}_2))\right]$$

$$\times \mathrm{E}_{H_0}\left[\left\{\cos(\pi \hat{p}_1(i,j;\hat{\boldsymbol{\theta}}_1)) - \cos(\pi \hat{p}_1(i,j;\boldsymbol{\theta}_2))\right\}^2\right],$$

where the last statement follows from logic similar to that used in Theorem 2.2.2. Then using the fact that the integrand in the second expectation is bounded, applying Lemma A.0.6 and the Lebesgue Dominated Convergence Theorem yields the result.

CHAPTER 3
PITMAN ASYMPTOTIC RELATIVE EFFICIENCIES

3.1 Introduction

To compute Pitman ARE's, a model of dependence must be adopted to serve as an alternative to the null hypothesis of independence. Konijn (1954, p. 300) states that "the crucial point is the specification of a class of alternatives which is (i) sufficiently wide to include some approximation to any situation that may arise in this class of problems, and (ii) manageable mathematically." For tests involving a change in location of a distribution, shift alternatives form a satisfactory idealization to a wide class of problems and are quite amenable to mathematical analysis. But because of the innumerable ways dependence can manifest itself, our situation cannot be expected to lend itself as easily to so simple a model. This is not necessarily a fatal blow, however, since when considering a model in the context of local alternatives, there is reason to believe the specific form of the model is of little consequence. Witness the agreement in the Pitman ARE's reported in Section 1.1 for the bivariate case when several different models were used. Thus, although we propose a model that is intuitively appealing in some aspects, our main reason for choosing it is for its mathematical tractability. We require that the model be a function of a nonnegative real-valued parameter Δ such that as $\Delta \to 0$, the sequence of alternatives defined by this model will converge to the null hypothesis. In fact, it is necessary that the convergence of this sequence of alternatives occurs at such a rate so that it is contiguous to the null hypothesis. This necessity is two-fold. In doing calculations under the sequence of alternatives, contiguity allows us to use an approximating quantity in finding the

limiting distribution of statistics of interest and also aids in determining the form of that limiting distribution.

3.2 Model 1

A generalization of the model apparently first studied by Konijn (1954) is given by

$$\begin{aligned} \boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}^{(1)} \\ \boldsymbol{X}^{(2)} \end{pmatrix} &= \begin{pmatrix} (1-\Delta)\boldsymbol{Y}^{(1)} + \Delta \boldsymbol{M}_1 \boldsymbol{Y}^{(2)} \\ \Delta \boldsymbol{M}_2 \boldsymbol{Y}^{(1)} + (1-\Delta)\boldsymbol{Y}^{(2)} \end{pmatrix} \\ &= \begin{pmatrix} (1-\Delta)\boldsymbol{I}_{r_1} & \Delta \boldsymbol{M}_1 \\ \Delta \boldsymbol{M}_2 & (1-\Delta)\boldsymbol{I}_{r_2} \end{pmatrix} \begin{pmatrix} \boldsymbol{Y}^{(1)} \\ \boldsymbol{Y}^{(2)} \end{pmatrix} \qquad (3.1) \\ &= \boldsymbol{A}_\Delta \begin{pmatrix} \boldsymbol{Y}^{(1)} \\ \boldsymbol{Y}^{(2)} \end{pmatrix} = \boldsymbol{A}_\Delta \boldsymbol{Y}, \end{aligned}$$

where $\boldsymbol{Y}^{(1)}$ and $\boldsymbol{Y}^{(2)}$ are independent random vectors that are $r_1 \times 1$ and $r_2 \times 1$, respectively, $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$ are arbitrary (known) matrices of dimensions $r_1 \times r_2$ and $r_2 \times r_1$, respectively, and $0 \leq \Delta \leq 1/2$. Notice that for $r_1 = r_2 \equiv r$ and $\boldsymbol{M}_1 = \boldsymbol{M}_2 = \boldsymbol{I}_r$, $\Delta = 1/2$ implies $\boldsymbol{X}^{(1)} = \boldsymbol{X}^{(2)}$ (perfect correlation), while $\Delta = 0$ corresponds to the null hypothesis of independence. Thus we can restate the testing problem as $H_0 : \Delta = 0$ vs. $H_1 : 0 < \Delta \leq 1/2$. Since $\boldsymbol{Y} = \boldsymbol{A}_\Delta^{-1}\boldsymbol{X}$ is a nonsingular linear transformation, the density function of $\boldsymbol{X}$ can be expressed as $f_X(\boldsymbol{x};\Delta) = \text{abs}\left(|\boldsymbol{A}_\Delta|^{-1}\right) f_Y\left(\boldsymbol{A}_\Delta^{-1}\boldsymbol{x}\right)$, where $f_Y(\boldsymbol{y}) = f_1\left(\boldsymbol{y}^{(1)}\right) f_2\left(\boldsymbol{y}^{(2)}\right)$ is the density function of $\boldsymbol{Y}$. We assume that the distributions of both $\boldsymbol{Y}^{(1)}$ and $\boldsymbol{Y}^{(2)}$ are elliptically symmetric with dispersion parameters Σ_1 and Σ_2, respectively, and are centered at $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$, respectively. In other words, $f_k\left(\boldsymbol{x}^{(k)}\right) = C_k g_k\left((\boldsymbol{x}^{(k)} - \boldsymbol{\theta}_k)'\Sigma_k^{-1}(\boldsymbol{x}^{(k)} - \boldsymbol{\theta}_k)\right)$, where Σ_k is a positive-definite matrix and $g_k()$ does not depend on $\boldsymbol{\theta}_k$ or Σ_k.

Such a model might conceivably arise when considering a battery of psychological or psychophysical tests administered to a group of subjects, with the goal of classifying the outcomes relative to certain independent "factors." Suppose that *apparently* the outcomes of one set of tests are practically determined by one factor, and the outcomes of another set of tests are practically determined by a second independent factor. In order to test this hypothesis, Model 1 could be used, since the alternative might reasonably be that all the outcomes depend, to varying degrees, on both factors.

3.2.1 Contiguity

We wish to show that the sequence of alternatives $H_1 : \Delta_n = n^{-1/2}\Delta_0$, where $\Delta_0 > 0$, is contiguous to the null hypothesis. To achieve this we follow the rationale of Hájek and Šidák (1967, pp. 201-214), which we outline here. Let $L(\boldsymbol{x}; \Delta_n) = f_X(\boldsymbol{x}; \Delta_n)/f_X(\boldsymbol{x}; 0)$, and $\Lambda_n = \log \prod_{i=1}^n L(\boldsymbol{X}_i; \Delta_n) = \sum_{i=1}^n \log L(\boldsymbol{X}_i; \Delta_n)$. LeCam's first lemma asserts that if $\Lambda_n \sim AN(-\sigma^2/2, \sigma^2)$, then the densities $\prod_{i=1}^n f_X(\boldsymbol{x}_i; \Delta_n)$ are contiguous to the densities $\prod_{i=1}^n f_X(\boldsymbol{x}_i; 0)$. Another way of expressing this is to say that the sequence of alternatives Δ_n is contiguous to the null hypothesis ($\Delta_0 = 0$). If $W_n = 2\sum_{i=1}^n [L(\boldsymbol{X}_i; \Delta_n)^{1/2} - 1]$, LeCam's second lemma states that contiguity will follow if, under H_0, the summands $\log L(\boldsymbol{X}_i; \Delta_n)$ are uniformly asymptotically negligible (UAN) and $W_n \sim AN(-\sigma^2/4, \sigma^2)$. Because the summands depend on n, finding the limiting distribution of W_n directly is quite difficult so they consider the first-order approximation of W_n. This approximation is expressed as $T_n = \Delta_n \sum_{i=1}^n L'(\boldsymbol{X}_i; 0)$, where $L'(\boldsymbol{x}; 0) \equiv \frac{\partial}{\partial \Delta} L(\boldsymbol{x}; \Delta)|_{\Delta=0}$. Hájek and Šidák demonstrate the contiguity for a univariate shift alternative (pp. 210–213) and a univariate scale alternative (pp. 213–214) by showing that under the null hypothesis, $W_n = T_n - \sigma^2/4 + o_p(1)$, the UAN condition holds, and $T_n \sim AN(0, \sigma^2)$. Randles (1989) has extended the arguments to show contiguity for a multivariate shift alternative. Noting that Model 1

is a multivariate extension of a scale alternative leads us to emulate the methods of Randles in extending the proof of contiguity. A sketch of this extension is included in Appendix B. Since we will need the form of T_n in determining limiting distributions under Δ_n, we prove the asymptotic normality of T_n presently. Considering previous discussion regarding the invariance of the statistics $-n \log V$ and $\hat{Q}_n$, we assume hereafter that $\boldsymbol{\theta}_k = \mathbf{0}$ and $\boldsymbol{\Sigma}_k = \boldsymbol{I}_{r_k}$. We discuss the ramifications of this assumption with respect to $-n \log S^{J_0}$ later. First we need to find the expression for $L'(\boldsymbol{x}; 0)$.

<u>Lemma 3.2.1</u> For Model 1 (given in (3.1)),

$$
\begin{aligned}
L'(\boldsymbol{x};0) = {} & 2\left(\boldsymbol{x}^{(1)\prime}\boldsymbol{x}^{(1)}\phi_1(\boldsymbol{x}^{(1)\prime}\boldsymbol{x}^{(1)}) + \frac{r_1}{2}\right) \\
& + 2\left(\boldsymbol{x}^{(2)\prime}\boldsymbol{x}^{(2)}\phi_2(\boldsymbol{x}^{(2)\prime}\boldsymbol{x}^{(2)}) + \frac{r_2}{2}\right) \\
& - 2\boldsymbol{x}^{(1)\prime}\left(\phi_1(\boldsymbol{x}^{(1)\prime}\boldsymbol{x}^{(1)})\boldsymbol{M}_1 + \phi_2(\boldsymbol{x}^{(2)\prime}\boldsymbol{x}^{(2)})\boldsymbol{M}_2'\right)\boldsymbol{x}^{(2)},
\end{aligned}
$$

where $\phi_k(t) = g_k'(t)/g_k(t)$.

<u>Proof of Lemma 3.2.1</u> Since

$$
L(\boldsymbol{x};\Delta) = \mathrm{abs}\left(|\boldsymbol{A}_\Delta|^{-1}\right)\frac{f_Y\left(\boldsymbol{A}_\Delta^{-1}\boldsymbol{x}\right)}{f_Y(\boldsymbol{x})},
$$

we see that

$$
L'(\boldsymbol{x};\Delta) = \left(\frac{\partial}{\partial\Delta}\mathrm{abs}\left(|\boldsymbol{A}_\Delta|^{-1}\right)\right)\frac{f_Y\left(\boldsymbol{A}_\Delta^{-1}\boldsymbol{x}\right)}{f_Y(\boldsymbol{x})} + \mathrm{abs}\left(|\boldsymbol{A}_\Delta|^{-1}\right)\left(\frac{\partial}{\partial\Delta}\frac{f_Y\left(\boldsymbol{A}_\Delta^{-1}\boldsymbol{x}\right)}{f_Y(\boldsymbol{x})}\right).
$$

Now

$$
\frac{\partial}{\partial\Delta}\mathrm{abs}\left(|\boldsymbol{A}_\Delta|^{-1}\right) = -\mathrm{abs}\left(|\boldsymbol{A}_\Delta|^{-1}\right)\mathrm{tr}\left(-\boldsymbol{A}_\Delta^{-1}\boldsymbol{P}\right),
$$

where

$$
\boldsymbol{P} \equiv \frac{\partial}{\partial\Delta}\boldsymbol{A}_\Delta = \begin{pmatrix} \boldsymbol{I}_{r_1} & -\boldsymbol{M}_1 \\ -\boldsymbol{M}_2 & \boldsymbol{I}_{r_2} \end{pmatrix},
$$

so that

$$\left.\frac{\partial}{\partial\Delta}\mathrm{abs}\left(|\boldsymbol{A}_\Delta|^{-1}\right)\right|_{\Delta=0} = \mathrm{abs}\left(|\boldsymbol{A}_0|^{-1}\right)\mathrm{tr}\left(\boldsymbol{A}_0^{-1}\boldsymbol{P}\right)$$

$$= \mathrm{abs}\left(|\boldsymbol{I}_{r_1+r_2}|^{-1}\right)\mathrm{tr}\left(\boldsymbol{I}_{r_1+r_2}^{-1}\boldsymbol{P}\right)$$

$$= \mathrm{tr}(\boldsymbol{P}) = r_1 + r_2.$$

Also

$$\frac{\partial}{\partial\Delta}\frac{f_Y\left(\boldsymbol{A}_\Delta^{-1}\boldsymbol{x}\right)}{f_Y(\boldsymbol{x})} = \left(\frac{\partial}{\partial\Delta}\boldsymbol{A}_\Delta^{-1}\boldsymbol{x}\right)'\frac{\nabla f_Y(\boldsymbol{x})}{f_Y(\boldsymbol{x})}$$

$$= \left(\boldsymbol{A}_\Delta^{-1}\boldsymbol{P}\boldsymbol{A}_\Delta^{-1}\boldsymbol{x}\right)'\frac{\nabla f_Y(\boldsymbol{x})}{f_Y(\boldsymbol{x})},$$

so that

$$\left.\frac{f_Y\left(\boldsymbol{A}_\Delta^{-1}\boldsymbol{x}\right)}{f_Y(\boldsymbol{x})}\right|_{\Delta=0} = \left(\boldsymbol{A}_0^{-1}\boldsymbol{P}\boldsymbol{A}_0^{-1}\boldsymbol{x}\right)'\frac{\nabla f_Y(\boldsymbol{x})}{f_Y(\boldsymbol{x})}$$

$$= \left(\boldsymbol{I}_{r_1+r_2}^{-1}\boldsymbol{P}\boldsymbol{I}_{r_1+r_2}^{-1}\boldsymbol{x}\right)'\frac{\nabla f_Y(\boldsymbol{x})}{f_Y(\boldsymbol{x})}$$

$$= (\boldsymbol{P}\boldsymbol{x})'\frac{\nabla f_Y(\boldsymbol{x})}{f_Y(\boldsymbol{x})}$$

$$= \left(\begin{pmatrix}\boldsymbol{I}_{r_1} & -\boldsymbol{M}_1\\ -\boldsymbol{M}_2 & \boldsymbol{I}_{r_2}\end{pmatrix}\begin{pmatrix}\boldsymbol{x}^{(1)}\\ \boldsymbol{x}^{(2)}\end{pmatrix}\right)'\begin{pmatrix}\dfrac{\nabla f_1\left(\boldsymbol{x}^{(1)}\right)}{f_1\left(\boldsymbol{x}^{(1)}\right)}\\ \dfrac{\nabla f_2\left(\boldsymbol{x}^{(2)}\right)}{f_2\left(\boldsymbol{x}^{(2)}\right)}\end{pmatrix}$$

$$= 2(\boldsymbol{x}^{(1)} - \boldsymbol{M}_1\boldsymbol{x}^{(2)})'\boldsymbol{x}^{(1)}\phi_1(\boldsymbol{x}^{(1)\prime}\boldsymbol{x}^{(1)})$$

$$+ 2(-\boldsymbol{M}_2\boldsymbol{x}^{(1)} + \boldsymbol{x}^{(2)})'\boldsymbol{x}^{(2)}\phi_2(\boldsymbol{x}^{(2)\prime}\boldsymbol{x}^{(2)}),$$

because

$$\frac{\nabla f_k\left(\boldsymbol{x}^{(k)}\right)}{f_k\left(\boldsymbol{x}^{(k)}\right)} = 2\boldsymbol{x}^{(k)}\,\frac{g_k'\left(\boldsymbol{x}^{(k)\prime}\boldsymbol{x}^{(k)}\right)}{g_k\left(\boldsymbol{x}^{(k)\prime}\boldsymbol{x}^{(k)}\right)}.$$

The result follows immediately. □

Recall that we can represent $\boldsymbol{X}^{(k)}$ as $R^{(k)}\boldsymbol{U}^{(k)}$ when $\boldsymbol{X}^{(k)}$ has a spherically symmetric distribution (see Section 2.2). Using this form for $\boldsymbol{X}^{(k)}$ gives

$$T_n = n^{-1/2}\sum_{i=1}^{n} 2\Delta_0\Big\{\Big((R_i^{(1)})^2\phi_1((R_i^{(1)})^2) + \frac{r_1}{2}\Big) + \Big((R_i^{(2)})^2\phi_2((R_i^{(2)})^2) + \frac{r_2}{2}\Big) - \boldsymbol{U}_i^{(1)\prime}\boldsymbol{R}_i\boldsymbol{U}_i^{(2)}\Big\},$$

where

$$\boldsymbol{R}_i = R_i^{(1)}R_i^{(2)}\left(\phi_1((R_i^{(1)})^2)\boldsymbol{M}_1 + \phi_2((R_i^{(2)})^2)\boldsymbol{M}_2'\right).$$

<u>Lemma 3.2.2</u> If $\mathrm{E}_{H_0}\left[(R^{(k)})^4\phi_k^2((R^{(k)})^2)\right] < \infty$, $\mathrm{E}_{H_0}\left[(R^{(k)})^2\phi_k^2((R^{(k)})^2)\right] < \infty$, and $\mathrm{E}_{H_0}\left[(R^{(k)})^2\right] < \infty$, then $\mathrm{V}_{H_0}[T_n] \equiv \sigma^2 < \infty$ and $T_n \sim AN(0, \sigma^2)$.

<u>Proof of Lemma 3.2.2</u> To guarantee that $\sigma^2 < \infty$, it is sufficient for

$$\mathrm{E}_{H_0}\left[\left\{\Big((R^{(1)})^2\phi_1((R^{(1)})^2) + \frac{r_1}{2}\Big) + \Big((R^{(2)})^2\phi_2((R^{(2)})^2) + \frac{r_2}{2}\Big) - \boldsymbol{U}^{(1)\prime}\boldsymbol{R}\boldsymbol{U}^{(2)}\right\}^2\right] < \infty,$$

where

$$\boldsymbol{R} = R^{(1)}R^{(2)}\left(\phi_1((R^{(1)})^2)\boldsymbol{M}_1 + \phi_2((R^{(2)})^2)\boldsymbol{M}_2'\right).$$

Thus $\sigma^2 < \infty$ if $\mathrm{E}_{H_0}\left[(R^{(k)})^4\phi_k^2((R^{(k)})^2)\right] < \infty$ and $\mathrm{E}_{H_0}\left[\left(\boldsymbol{U}^{(1)\prime}\boldsymbol{R}\boldsymbol{U}^{(2)}\right)^2\right] < \infty$, where the second expectation is easily seen to be finite if $\mathrm{E}_{H_0}\left[(R^{(k)})^2\phi_k^2((R^{(k)})^2)\right] < \infty$ and $\mathrm{E}_{H_0}\left[(R^{(k)})^2\right] < \infty$. Appendix B shows further that if $\sigma^2 < \infty$, then $\mathrm{E}_{H_0}[T_n] = 0$. Since the terms are iid, an application of the central limit theorem gives the result. □

Thus, we have established conditions under which contiguity holds. In the next section we work out the limiting distributions of $\hat{Q}_n$, $-n\log V$, and $-n\log S^{J_0}$.

3.2.2 Limiting Distributions Under Δ_n

LeCam's third lemma states that if, under H_0, $\binom{S_n}{\Lambda_n} \sim AN\left(\binom{\mu_1}{-\sigma_2^2/2}, \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}\right)$, then $S_n \sim AN(\mu_1 + \sigma_{12}, \sigma_1^2)$ under a contiguous sequence of alternatives. Finding the limiting distributions of $\hat{Q}_n$, $-n \log V$, and $-n \log S^{J_0}$ under the the contiguous sequence of alternatives Δ_n will involve showing that under H_0, $\binom{S_n}{T_n} \sim AN\left(\binom{0}{0}, \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}\right)$, where S_n is an appropriately defined statistic. Then, under Δ_n, $S_n \sim AN(\sigma_{12}, \sigma_1^2)$ and from this it will be possible to determine the limiting distribution of the statistic of interest. We begin by finding the asymptotic distribution for $\hat{Q}_n$. We assume throughout the rest of the chapter that the moment conditions in Lemma 3.2.2 are satisfied.

Theorem 3.2.1 Under Δ_n,

$$\hat{Q}_n \xrightarrow{d} \chi^2_{r_1 r_2}\left(\frac{4\Delta_0}{r_1 r_2} \text{vec}\left(\mathrm{E}_{H_0}[\boldsymbol{R}]\right)' \text{vec}\left(\mathrm{E}_{H_0}[\boldsymbol{R}]\right)\right).$$

Proof of Theorem 3.2.1 Let $\boldsymbol{a} = (a_1, a_2)'$ be an arbitrary pair of constants not both zero. If $S_n = n^{-1/2} \sum_{i=1}^n \boldsymbol{U}_i^{(1)'}(\sqrt{r_1 r_2}\boldsymbol{B})\boldsymbol{U}_i^{(2)}$ (cf. Theorem 2.2.1), then

$$\boldsymbol{a}' \begin{pmatrix} S_n \\ T_n \end{pmatrix} =$$

$$n^{-1/2} \sum_{i=1}^n 2\Delta_0 \Big\{ a_2 \left((R_i^{(1)})^2 \phi_1((R_i^{(1)})^2) + \frac{r_1}{2} \right) + a_2 \left((R_i^{(2)})^2 \phi_2((R_i^{(2)})^2) + \frac{r_2}{2} \right)$$

$$+ \boldsymbol{U}_i^{(1)'} \left(a_2 \boldsymbol{R}_i + a_1 \frac{\sqrt{r_1 r_2}}{2\Delta_0} \boldsymbol{B} \right) \boldsymbol{U}_i^{(2)} \Big\}.$$

Since $\mathrm{E}_{H_0}[S_n]$ and $\mathrm{E}_{H_0}[T_n]$ are zero, $\mathrm{E}_{H_0}[\boldsymbol{a}'(S_n, T_n)'] = 0$. Also, the summands are iid where the three terms in each summand are uncorrelated with mean zero so that,

$$\begin{aligned}
\mathrm{V}_{H_0}[\boldsymbol{a}'(S_n, T_n)] &= a_1^2 \mathrm{V}_{H_0}[S_n] + a_2^2 \mathrm{V}_{H_0}[T_n] + 2a_1a_2 \mathrm{Cov}_{H_0}[S_n, T_n] \\
&= a_1^2 \mathrm{vec}(\boldsymbol{B})' \mathrm{vec}(\boldsymbol{B}) + a_2^2\sigma^2 \\
&\quad + 2a_1a_2 \cdot 2\Delta_0\sqrt{r_1r_2}\mathrm{E}_{H_0}\left[\boldsymbol{U}^{(1)\prime}\boldsymbol{R}\boldsymbol{U}^{(2)}\boldsymbol{U}^{(1)\prime}\boldsymbol{B}\boldsymbol{U}^{(2)}\right].
\end{aligned}$$

Further, using (2.2),

$$\begin{aligned}
\mathrm{E}_{H_0}\left[\boldsymbol{U}^{(1)\prime}\boldsymbol{R}\boldsymbol{U}^{(2)}\boldsymbol{U}^{(1)\prime}\boldsymbol{B}\boldsymbol{U}^{(2)}\right] &= \frac{1}{r_2}\mathrm{E}_{H_0}\left[\boldsymbol{U}^{(1)\prime}\boldsymbol{R}\boldsymbol{B}'\boldsymbol{U}^{(1)}\right] \\
&= \frac{1}{r_1r_2}\mathrm{vec}(\boldsymbol{B})' \mathrm{vec}\left(\mathrm{E}_{H_0}[\boldsymbol{R}]\right),
\end{aligned}$$

so that $\sigma_1^2 = \mathrm{vec}(\boldsymbol{B})' \mathrm{vec}(\boldsymbol{B})$, $\sigma_2^2 = \sigma^2$, and $\sigma_{12} = \mathrm{vec}(\boldsymbol{B})' \left(\frac{2\Delta_0}{\sqrt{r_1r_2}}\mathrm{vec}\left(\mathrm{E}_{H_0}[\boldsymbol{R}]\right)\right)$. The asymptotic normality of $\boldsymbol{a}'(S_n, T_n)'$ follows by applying the central limit theorem. Then under Δ_n,

$$S_n \sim AN\left(\mathrm{vec}(\boldsymbol{B})' \left(\frac{2\Delta_0}{\sqrt{r_1r_2}}\mathrm{vec}\left(\mathrm{E}_{H_0}[\boldsymbol{R}]\right)\right), \mathrm{vec}(\boldsymbol{B})' \mathrm{vec}(\boldsymbol{B})\right),$$

with Lemmas 2.2.1 and 2.2.2 giving

$$Q_n \xrightarrow{d} \chi^2_{r_1r_2}\left(\frac{4\Delta_0^2}{r_1r_2}\mathrm{vec}\left(\mathrm{E}_{H_0}[\boldsymbol{R}]\right)' \mathrm{vec}\left(\mathrm{E}_{H_0}[\boldsymbol{R}]\right)\right).$$

The result follows by noting that since the difference between Q_n and $\hat{Q}_n$ converges in probability to zero under the null hypothesis (see the proof of Theorem 2.2.2), by definition, the same difference will converge in probability to zero under any contiguous sequence of alternatives. Thus Q_n and $\hat{Q}_n$ will have the same limiting distribution under Δ_n. $\square$

We next find the asymptotic distribution of $-n \log V$ under Δ_n. It is easy to show, using U-statistics, that if $\mathrm{E}_{H_0}[X_{il}^4] < \infty$, then $-n \log V \xrightarrow{d} \chi^2_{r_1r_2}$ under H_0. To

determine the limiting distribution under Δ_n, we need to find a simple approximating quantity. Puri and Sen (1971, p. 364) show that

$$\left| -n \log V - n \sum_{s=1}^{r_1} \sum_{s'=1}^{r_1} \sum_{t=1}^{r_2} \sum_{t'=1}^{r_2} \hat{\rho}_{st} \hat{\rho}_{s't'} \hat{\rho}_{ss'}^{(1)} \hat{\rho}_{tt'}^{(2)} \right| = O_p(n^{-1})$$

where $\hat{\boldsymbol{\rho}}$, the sample correlation matrix, is partitioned into $\hat{\boldsymbol{\rho}}_{11}^{-1} = \{\hat{\rho}_{ss'}^{(1)}\}_{r_1 \times r_1}$, $\hat{\boldsymbol{\rho}}_{22}^{-1} = \{\hat{\rho}_{tt'}^{(2)}\}_{r_2 \times r_2}$, and $\hat{\boldsymbol{\rho}}_{12} = \{\hat{\rho}_{st}\}_{r_1 \times r_2} = \hat{\boldsymbol{\rho}}_{21}'$. Since $\hat{\boldsymbol{\rho}}_{kk} = \boldsymbol{I}_{r_k} + \boldsymbol{o}_p(1)$, we use Slutsky's Theorem to achieve the following simplification,

$$\begin{aligned} -n \log V &= n \operatorname{vec}(\hat{\boldsymbol{\rho}}_{12})' (\hat{\boldsymbol{\rho}}_{11}^{-1} \otimes \hat{\boldsymbol{\rho}}_{22}^{-1}) \operatorname{vec}(\hat{\boldsymbol{\rho}}_{12}) + O_p(n^{-1}) \\ &= n \operatorname{vec}(\hat{\boldsymbol{\rho}}_{12})' \operatorname{vec}(\hat{\boldsymbol{\rho}}_{12}) + o_p(1). \end{aligned}$$

Then, since $\hat{\boldsymbol{\rho}}_{12} = n^{-1} \boldsymbol{A}_{12}$ and $n^{-1} \sum_{i=1}^{n} \boldsymbol{X}_i^{(k)} = \boldsymbol{o}_p(1)$, we use Slutsky's theorem again to get

$$\begin{aligned} -n \log V &= n^{-1} \operatorname{vec}(\boldsymbol{A}_{12})' \operatorname{vec}(\boldsymbol{A}_{12}) + o_p(1) \\ &= n^{-1} \operatorname{vec}\left(\sum_{i=1}^{n} \boldsymbol{X}_i^{(1)} \boldsymbol{X}_i^{(2)'}\right)' \operatorname{vec}\left(\sum_{i=1}^{n} \boldsymbol{X}_i^{(1)} \boldsymbol{X}_i^{(2)'}\right) + o_p(1) \\ &= n^{-1} \sum_{i=1}^{n} \sum_{j=1}^{n} \operatorname{tr}\left(\left(\boldsymbol{X}_i^{(1)} \boldsymbol{X}_i^{(2)'}\right)' \boldsymbol{X}_j^{(1)} \boldsymbol{X}_j^{(2)'}\right) + o_p(1) \\ &= n^{-1} \sum_{i=1}^{n} \sum_{j=1}^{n} \boldsymbol{X}_i^{(1)'} \boldsymbol{X}_j^{(1)} \boldsymbol{X}_i^{(2)'} \boldsymbol{X}_j^{(2)} + o_p(1) \\ &= n^{-1} \sum_{i=1}^{n} \sum_{j=1}^{n} \left(R_i^{(1)} R_j^{(1)} R_i^{(2)} R_j^{(2)}\right) \boldsymbol{U}_i^{(1)'} \boldsymbol{U}_j^{(1)} \boldsymbol{U}_i^{(2)'} \boldsymbol{U}_j^{(2)} + o_p(1). \end{aligned} \tag{3.2}$$

Thus we have a convenient approximating quantity for $-n \log V$. In fact, comparing the approximations for $\hat{Q}_n$ in (2.1) and $-n \log V$ in (3.2) reveals the underlying similarity in structure between these two statistics, which at first is not readily apparent. To find the limiting distribution of $-n \log V$ under Δ_n, we proceed as in the proof of Theorem 3.2.1.

Theorem 3.2.2 Under Δ_n,

$$-n \log V \xrightarrow{d} \chi^2_{r_1 r_2}\left(\frac{4\Delta_0}{r_1^2 r_2^2}\mathrm{vec}\left(\mathrm{E}_{H_0}\left[R^{(1)}R^{(2)}\boldsymbol{R}\right]\right)'\mathrm{vec}\left(\mathrm{E}_{H_0}\left[R^{(1)}R^{(2)}\boldsymbol{R}\right]\right)\right).$$

Proof of Theorem 3.2.2 Let

$$\begin{aligned} S_n &= n^{-1/2}\mathrm{vec}\,(\boldsymbol{B})'\,\mathrm{vec}\left(\sum_{i=1}^{n}\boldsymbol{X}_i^{(1)}\boldsymbol{X}_i^{(2)'}\right) \\ &= n^{-1/2}\sum_{i=1}^{n}\boldsymbol{X}_i^{(1)'}\boldsymbol{B}\boldsymbol{X}_i^{(2)} \\ &= n^{-1/2}\sum_{i=1}^{n}\boldsymbol{U}_i^{(1)'}(R_i^{(1)}R_i^{(2)}\boldsymbol{B})\boldsymbol{U}_i^{(2)}. \end{aligned}$$

The key quantity is again

$$\begin{aligned} \sigma_{12} &= \mathrm{Cov}_{H_0}\left[S_n, T_n\right] \\ &= 2\Delta_0\mathrm{E}_{H_0}\left[\boldsymbol{U}^{(1)'}\boldsymbol{R}\boldsymbol{U}^{(2)'}\boldsymbol{U}^{(1)}R^{(1)}R^{(2)}\boldsymbol{B}\boldsymbol{U}^{(2)}\right] \\ &= \mathrm{vec}\,(\boldsymbol{B})'\left(\frac{2\Delta_0}{r_1 r_2}\mathrm{vec}\left(\mathrm{E}_{H_0}\left[R^{(1)}R^{(2)}\boldsymbol{R}\right]\right)\right). \end{aligned}$$

Applying Lemma 2.2.1 and then Lemma 2.2.2 yields the result. □

If we adopt the notations $\mathrm{sgn}\,(\boldsymbol{x}) = (\mathrm{sgn}\,(x_1),\ldots,\mathrm{sgn}\,(x_n))'$ and $\tilde{\boldsymbol{x}} = (\tilde{x}_1,\ldots,\tilde{x}_n)'$, then use of the score function J_0 defined by (1.2) enables us to represent the matrix $\boldsymbol{T}_{12}$ described in Section 1.2.2 as

$$\boldsymbol{T}_{12} = n^{-1}\sum_{i=1}^{n}\mathrm{sgn}\left(\boldsymbol{X}_i^{(1)} - \widetilde{\boldsymbol{X}}^{(1)}\right)\mathrm{sgn}\left(\boldsymbol{X}_i^{(2)} - \widetilde{\boldsymbol{X}}^{(2)}\right)'.$$

Further, using Theorem 2.13 in Randles (1982) on the components of $\boldsymbol{T}_{12}$, we are free to replace the sample medians by the population medians (which we have assumed are zero) when considering the asymptotic distribution of $\boldsymbol{T}_{12}$. Continuing, we now have

$$\begin{aligned} \boldsymbol{T}_{12} &= n^{-1}\sum_{i=1}^{n}\mathrm{sgn}\left(\boldsymbol{X}_i^{(1)}\right)\mathrm{sgn}\left(\boldsymbol{X}_i^{(2)}\right)' + o_p(1) \\ &= n^{-1}\sum_{i=1}^{n}\mathrm{sgn}\left(\boldsymbol{U}_i^{(1)}\right)\mathrm{sgn}\left(\boldsymbol{U}_i^{(2)}\right)' + o_p(1). \end{aligned}$$

Combining this with the result that $-n \log S^{J_0} = n\mathrm{vec}\,(\boldsymbol{T}_{12})' \,\mathrm{vec}\,(\boldsymbol{T}_{12}) + O_p(n^{-1})$ (see Puri & Sen, 1971, p. 359), shows that

$$
\begin{aligned}
&- n \log S^{J_0} \\
&= n^{-1} \mathrm{vec} \left(\sum_{i=1}^{n} \mathrm{sgn}\left(\boldsymbol{U}_i^{(1)}\right) \mathrm{sgn}\left(\boldsymbol{U}_i^{(2)}\right)' \right)' \mathrm{vec} \left(\sum_{i=1}^{n} \mathrm{sgn}\left(\boldsymbol{U}_i^{(1)}\right) \mathrm{sgn}\left(\boldsymbol{U}_i^{(2)}\right)' \right) + o_p(1) \\
&= n^{-1} \sum_{i=1}^{n} \sum_{j=1}^{n} \mathrm{sgn}\left(\boldsymbol{U}_i^{(1)}\right)' \mathrm{sgn}\left(\boldsymbol{U}_j^{(1)}\right) \mathrm{sgn}\left(\boldsymbol{U}_i^{(2)}\right)' \mathrm{sgn}\left(\boldsymbol{U}_j^{(2)}\right) + o_p(1),
\end{aligned}
$$

where again it is of interest to note how this approximation compares to those of $\hat{Q}_n$ in (2.1) and $-n \log V$ in (3.2).

On a cautionary note, because of the noninvariance of $-n \log S^{J_0}$, all subsequent derivations apply only when Σ_k is in reality a diagonal matrix. Recall with the other statistics there was no loss of generality because of their invariance under $\mathcal{G}$. This is probably not a serious concern, since as Randles (1989) has noted in the comparison of his (affine-invariant) interdirection sign test with a (nonaffine-invariant) component-wise sign test, spherical symmetry is favorable to the component-wise test. He further states that although it might be possible to improve slightly the efficiency of the component-wise test over certain points in the alternative, the result is a drastic depreciation of its efficiency over the rest of the alternative space. This generally is not a desirable property of a statistic. The same logic applies in the present situation. After deriving a moment needed in the subsequent theorem, we proceed exactly as in Theorem 3.2.1.

Lemma 3.2.3 If $\boldsymbol{U}^{(k)} \sim \mathrm{Uniform}(\Omega_{r_k})$, then

$$
\mathrm{E}\left[|U_t^{(k)}| \right] = \frac{\Gamma\left(\frac{r_k}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{r_k + 1}{2}\right)}.
$$

Proof of Lemma 3.2.3 A point on an r_k-dimensional unit sphere can be uniquely represented (provided $r_k \geq 2$) by $r_k - 1$ angles $\eta_1, \ldots, \eta_{r_k-1}$ and the equations

$$U_1^{(k)} = \sin \eta_1 \sin \eta_2 \ldots \sin \eta_{r_k-2} \sin \eta_{r_k-1}$$

$$U_2^{(k)} = \sin \eta_1 \sin \eta_2 \ldots \sin \eta_{r_k-2} \cos \eta_{r_k-1}$$

$$U_3^{(k)} = \sin \eta_1 \sin \eta_2 \ldots \sin \eta_{r_k-3} \cos \eta_{r_k-2}$$

$$U_4^{(k)} = \sin \eta_1 \sin \eta_2 \ldots \sin \eta_{r_k-4} \cos \eta_{r_k-3}$$

$$\vdots$$

$$U_{r_k-2}^{(k)} = \sin \eta_1 \sin \eta_2 \cos \eta_3$$

$$U_{r_k-1}^{(k)} = \sin \eta_1 \cos \eta_2$$

$$U_{r_k}^{(k)} = \cos \eta_1 .$$

If the joint density of the angles is proportional to $\sin^{r_k-2} \eta_1 \sin^{r_k-3} \eta_2 \cdots \sin \eta_{r_k-2}$, where $0 < \eta_i < \pi$, $i = 1, \ldots, r_k - 2$ and $0 < \eta_{r_k-1} < 2\pi$, then $\boldsymbol{U}^{(k)}$ will have the uniform distribution on the unit hypersphere of dimension r_k. Clearly the angles are independent, with η_{r_k-1} uniformly distributed on $(0, 2\pi)$ and the other angles having power sine densities on $(0, \pi)$. For instance, the marginal density for η_1 is

$$g(\eta_1) = \frac{\Gamma\left(\frac{r_k}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{r_k - 1}{2}\right)} \sin^{r_k-2} \eta_1 .$$

Thus,

$$\begin{aligned}
\mathrm{E}\left[|U_1^{(k)}|\right] &= \int_0^{\pi} |\cos\eta_1| g(\eta_1) d\eta_1 \\
&= \frac{2\Gamma\left(\frac{r_k}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{r_k-1}{2}\right)} \int_0^{\pi/2} \cos\eta_1 \sin^{r_k-2}\eta_1 d\eta_1 \\
&= \frac{2\Gamma\left(\frac{r_k}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{r_k-1}{2}\right)} \int_0^{1} z^{r_k-2} dz \\
&= \frac{2\Gamma\left(\frac{r_k}{2}\right)}{(r_k-1)\sqrt{\pi}\,\Gamma\left(\frac{r_k-1}{2}\right)} \\
&= \frac{\Gamma\left(\frac{r_k}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{r_k+1}{2}\right)}.
\end{aligned}$$

Since $|U_1^{(k)}| = 1$ when $r_k = 1$, the formula holds for $r_k \geq 1$. By symmetry, the result holds for $\mathrm{E}\left[|U_t^{(k)}|\right]$. □

Theorem 3.2.3 Under Δ_n,

$$-n \log S^{J_0} \xrightarrow{d} \chi^2_{r_1 r_2}\left(\frac{4\Delta_0^2 \Gamma^2\left(\frac{r_1}{2}\right)\Gamma^2\left(\frac{r_2}{2}\right)}{\pi^2\Gamma^2\left(\frac{r_1+1}{2}\right)\Gamma^2\left(\frac{r_2+1}{2}\right)} \mathrm{vec}\left(\mathrm{E}_{H_0}[\boldsymbol{R}]\right)' \mathrm{vec}\left(\mathrm{E}_{H_0}[\boldsymbol{R}]\right)\right).$$

Proof of Theorem 3.2.3 Let $S_n = n^{-1/2}\sum_{i=1}^n \mathrm{sgn}\left(\boldsymbol{U}_i^{(1)}\right)' \boldsymbol{B}\mathrm{sgn}\left(\boldsymbol{U}_i^{(2)}\right)$ so that the key quantity is again

$$\begin{aligned}
\sigma_{12} &= \mathrm{Cov}_{H_0}[S_n, T_n] \\
&= 2\Delta_0 \mathrm{E}_{H_0}\left[\boldsymbol{U}^{(1)\prime}\boldsymbol{R}\boldsymbol{U}^{(2)}\mathrm{sgn}\left(\boldsymbol{U}^{(1)}\right)'\boldsymbol{B}\mathrm{sgn}\left(\boldsymbol{U}^{(2)}\right)\right] \\
&= \mathrm{vec}(\boldsymbol{B})'\left(2\Delta_0\mathrm{E}_{H_0}\left[|U_1^{(1)}|\right]\mathrm{E}_{H_0}\left[|U_1^{(2)}|\right]\mathrm{vec}\left(\mathrm{E}_{H_0}[\boldsymbol{R}]\right)\right)
\end{aligned}$$

and using Lemma 3.2.3,

$$= \mathrm{vec}\,(\boldsymbol{B})' \left(\frac{2\Delta_0 \Gamma\left(\frac{r_1}{2}\right)\Gamma\left(\frac{r_2}{2}\right)}{\pi\Gamma\left(\frac{r_1+1}{2}\right)\Gamma\left(\frac{r_2+1}{2}\right)} \mathrm{vec}\left(\mathrm{E}_{H_0}\left[\boldsymbol{R}\right]\right) \right),$$

with Lemma 2.2.1 and Lemma 2.2.2 giving the result. □

3.2.3 Comparison of Statistics

We are now in a position to find the expressions for ARE($\hat{Q}_n, -n\log V$) and ARE($\hat{Q}_n, -n\log S^{J_0}$). Since each of the statistics have limiting noncentral $\chi^2_{r_1 r_2}$ distributions under Δ_n, Hannan (1956) has shown that the Pitman ARE is the ratio of the noncentrality parameters. Referring to Theorem 3.2.1 and Theorem 3.2.2 we are able to report that

$$\mathrm{ARE}(\hat{Q}_n, -n\log V) = \frac{r_1 r_2 \mathrm{vec}\left(\mathrm{E}_{H_0}\left[\boldsymbol{R}\right]\right)' \mathrm{vec}\left(\mathrm{E}_{H_0}\left[\boldsymbol{R}\right]\right)}{\mathrm{vec}\left(\mathrm{E}_{H_0}\left[R^{(1)}R^{(2)}\boldsymbol{R}\right]\right)' \mathrm{vec}\left(\mathrm{E}_{H_0}\left[R^{(1)}R^{(2)}\boldsymbol{R}\right]\right)}$$

$$= \frac{4\mathrm{vec}\,(\varphi_1 \boldsymbol{M}_1 + \varphi_2 \boldsymbol{M}_2')' \,\mathrm{vec}\,(\varphi_1 \boldsymbol{M}_1 + \varphi_2 \boldsymbol{M}_2')}{r_1 r_2 \mathrm{vec}\,(\boldsymbol{M}_1 + \boldsymbol{M}_2')' \,\mathrm{vec}\,(\boldsymbol{M}_1 + \boldsymbol{M}_2')},$$

where

$$\varphi_1 = \mathrm{E}_{H_0}\left[R^{(2)}\right] \mathrm{E}_{H_0}\left[R^{(1)}\phi_1((R^{(1)})^2)\right]$$

and

$$\varphi_2 = \mathrm{E}_{H_0}\left[R^{(1)}\right] \mathrm{E}_{H_0}\left[R^{(2)}\phi_2((R^{(2)})^2)\right].$$

Note that for $\boldsymbol{M}_1 = \boldsymbol{M}_2'$,

$$\mathrm{ARE}(\hat{Q}_n, -n\log V) = \frac{(\varphi_1 + \varphi_2)^2}{r_1 r_2}. \tag{3.3}$$

Also, using Theorem 3.2.1 and Theorem 3.2.3 we see that

$$\mathrm{ARE}(\hat{Q}_n, -n\log S^{J_0}) = \frac{\pi^2\Gamma^2\left(\frac{r_1+1}{2}\right)\Gamma^2\left(\frac{r_2+1}{2}\right)}{r_1 r_2 \Gamma^2\left(\frac{r_1}{2}\right)\Gamma^2\left(\frac{r_2}{2}\right)},$$

where it is of interest to note that there is no dependence on g_1 and g_2 or the form of the matrices $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$. Of course, as noted earlier, this result holds only when Σ_k is diagonal. For general Σ_k, (i.e., elliptically symmetric distributions), the Pitman ARE will depend on the underlying covariance structure of the $\boldsymbol{X}^{(1)}$'s and $\boldsymbol{X}^{(2)}$'s. In the next sections, we compute ARE($\hat{Q}_n, -n\log V$) for various choices of g_1 and g_2.

Because the formulas for the Pitman ARE's are quite complex, a visual aid can help reveal some of their structure. Therefore, we provide graphs which illustrate the Pitman ARE's for the three statistics in various simplified situations. One simplification is that identical distributions were used for both $\boldsymbol{X}^{(1)}$ and $\boldsymbol{X}^{(2)}$. Recall that the Pitman ARE is the ratio of sample sizes needed for competing tests to maintain the same limiting power and size when converging to the null hypothesis. Thus, if T_n and S_n are two competing sequences of tests, then loosely speaking, ARE$(T_n, S_n) > (=, <) 1$ implies that T_n requires fewer (equal, more) observations to maintain about the same power as S_n, meaning T_n is more (equally as, less) efficient. Since normally ARE(T_n,S_n) is not necessarily bounded, we have chosen to plot the Pitman ARE's for the dimensions given by the axes labeled $r1$ and $r2$ using the transformation 1/(1+ARE(S_n,T_n)). This results in the surface lying between 0 and 1 for any possible value of ARE(S_n,T_n). Theoretically, this keeps the visual comparison of situations where the competing statistics relative performances change direction on equal footing. The surface is now loosely interpreted as the ratio of the sample size required for T_n to the sum of the sample sizes required for both T_n and S_n. Practically, this means that in all cases T_n is doing better if the surface is below 1/2 and worse if it's above 1/2. For example, in Figure 3.1, $\hat{Q}_n$ is in the second position and the surface is below 1/2 (except at $r1 = r2 = 1$ when $\hat{Q}_n$ and $-n\log S^{J_0}$ are asymptotically equivalent), so it is more efficient than $-n\log S^{J_0}$ when at least one of $r1$ or $r2$ is bigger than 1 and equally as efficient in the bivariate case. Similar

graphs of the Pitman ARE of $\hat{Q}_n$ with $-n\log V$ follow the sections in which specific distributions have been assumed, allowing for the evaluation of φ_1 and φ_2.

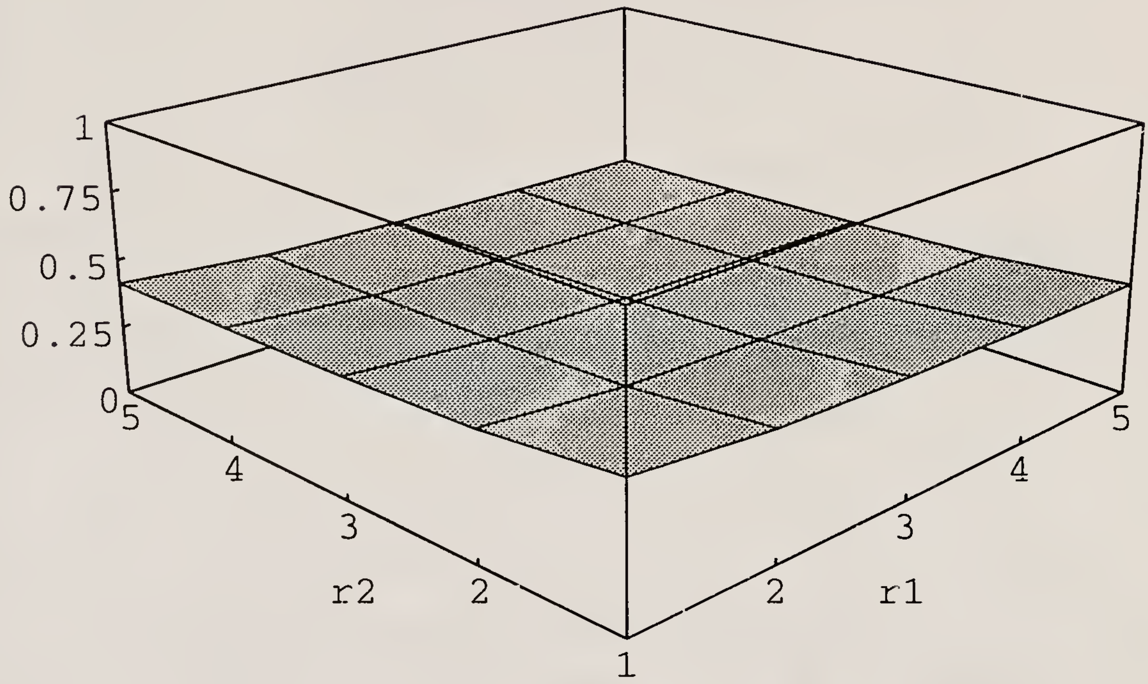

Figure 3.1. $1/(1+\mathrm{ARE}(-n\log S^{J_0}, \hat{Q}_n))$

Exponential power class

A convenient elliptically symmetric class of distributions is the exponential power class. The exponential parameter ν allows for a choice of distributions with varying

heaviness in their tails. This will enable us to evaluate $\mathrm{ARE}(\hat{Q}_n, -n\log V)$ when the underlying distributions of $\boldsymbol{X}^{(1)}$ and $\boldsymbol{X}^{(2)}$ have either heavy or light tails. When $\nu = 1$, this corresponds to the multivariate normal distribution. When $0 < \nu < 1$, the resulting distribution has heavier tails than the normal and when $\nu > 1$, the resulting distribution has lighter tails than the normal. In fact, as $\nu \to \infty$, the distribution becomes uniform.

Let

$$g_k(t) = \exp\left(-\frac{t^{\nu_k}}{c_k}\right),$$

and

$$C_k = \frac{\nu_k \Gamma\left(\frac{r_k}{2}\right)}{\Gamma\left(\frac{r_k}{2\nu_k}\right)}\left[\frac{\Gamma\left(\frac{r_k+2}{2\nu_k}\right)}{\pi r_k \Gamma\left(\frac{r_k}{2\nu_k}\right)}\right]^{r_k/2}, \qquad c_k = \left[\frac{r_k\Gamma\left(\frac{r_k}{2\nu_k}\right)}{\Gamma\left(\frac{r_k+2}{2\nu_k}\right)}\right]^{\nu_k} = d_k^{\nu_k}.$$

Then $\boldsymbol{X}^{(k)}$ has the exponential power distribution located at the origin with dispersion parameter $\boldsymbol{I}_{r_k}$ and exponential parameter ν_k ($\boldsymbol{X}^{(k)} \sim \mathrm{Exp}(\nu_k)$). The density function of $(R^{(k)})^2$ is given by:

$$h_k(t) = \frac{C_k \pi^{\nu_k/2}}{\Gamma\left(\frac{\nu_k}{2}\right)} t^{r_k/2-1} \exp\left(-\frac{t^{\nu_k}}{c_k}\right). \tag{3.4}$$

We first calculate a useful moment.

Lemma 3.2.4

$$\mathrm{E}_{H_0}\left[(R^{(k)})^a\right] = \frac{\Gamma\left(\frac{r_k+a}{2\nu_k}\right)}{\Gamma\left(\frac{r_k}{2\nu_k}\right)} d_k^{a/2} \qquad \text{provided } r_k + a > 0.$$

Proof of Lemma 3.2.4 Let T have the density h given by (3.4). Then

$$\mathrm{E}\left[T^{a/2}\right] = \frac{C_k \pi^{\nu_k/2}}{\Gamma\left(\frac{\nu_k}{2}\right)} \int_0^\infty t^{(r_k+a)/2-1} \exp(-(t/d_k)^{\nu_k}) dt$$

$$= \frac{C_k \pi^{\nu_k/2} d_k^{(r_k+a)/2}}{\Gamma\left(\frac{\nu_k}{2}\right)\nu_k} \int_0^\infty t^{(r_k+a)/(2\nu_k)-1} \exp(-t) dt$$

$$= \frac{C_k \pi^{\nu_k/2} d_k^{(r_k+a)/2} \Gamma\left(\frac{r_k+a}{2\nu_k}\right)}{\Gamma\left(\frac{\nu_k}{2}\right)\nu_k}$$

$$= \frac{\Gamma\left(\frac{r_k+a}{2\nu_k}\right)}{\Gamma\left(\frac{r_k}{2\nu_k}\right)} d_k^{a/2},$$

so long as $r_k + a > 0$. □

Theorem 3.2.4 If $\boldsymbol{X}^{(k)} \sim \mathrm{Exp}(\nu_k)$, with $\nu_k > -(r_k - 2)/4$, and we assume that $\boldsymbol{M}_1 = \boldsymbol{M}_2'$, then

$\mathrm{ARE}(\hat{Q}_n, -n \log V; \nu_1, \nu_2)$

$$= \frac{1}{r_1 r_2}\left(\nu_1 \left(\frac{d_2}{d_1}\right)^{1/2} \frac{\Gamma\left(\frac{r_2+1}{2\nu_2}\right)\Gamma\left(\frac{r_1+2\nu_1-1}{2\nu_1}\right)}{\Gamma\left(\frac{r_2}{2\nu_2}\right)\Gamma\left(\frac{r_1}{2\nu_1}\right)}\right.$$

$$\left. + \nu_2 \left(\frac{d_1}{d_2}\right)^{1/2} \frac{\Gamma\left(\frac{r_1+1}{2\nu_1}\right)\Gamma\left(\frac{r_2+2\nu_2-1}{2\nu_2}\right)}{\Gamma\left(\frac{r_1}{2\nu_1}\right)\Gamma\left(\frac{r_2}{2\nu_2}\right)}\right)^2.$$

Proof of Theorem 3.2.4 Since

$$\phi_k(t) = g_k'(t)/g_k(t) = -\frac{\nu_k}{c_k} t^{\nu_k - 1},$$

$$\varphi_1 = \mathrm{E}_{H_0}\left[R^{(2)}\right]\mathrm{E}_{H_0}\left[R^{(1)}\phi_1((R^{(1)})^2)\right]$$

$$= -\frac{\nu_1}{c_1}\mathrm{E}_{H_0}\left[R^{(2)}\right]\mathrm{E}_{H_0}\left[R^{(1)}R^{(1)^{2(\nu_1-1)}}\right]$$

$$= -\frac{\nu_1}{c_1}\mathrm{E}_{H_0}\left[R^{(2)}\right]\mathrm{E}_{H_0}\left[R^{(1)^{2\nu_1-1}}\right]$$

$$= -\frac{\nu_1}{c_1}\frac{\Gamma\left(\frac{r_2+1}{2\nu_2}\right)}{\Gamma\left(\frac{r_2}{2\nu_2}\right)}d_2^{1/2}\cdot\frac{\Gamma\left(\frac{r_1+2\nu_1-1}{2\nu_1}\right)}{\Gamma\left(\frac{r_1}{2\nu_1}\right)}d_1^{(2\nu_1-1)/2}$$

$$= -\nu_1\left(\frac{d_2}{d_1}\right)^{1/2}\frac{\Gamma\left(\frac{r_2+1}{2\nu_2}\right)\Gamma\left(\frac{r_1+2\nu_1-1}{2\nu_1}\right)}{\Gamma\left(\frac{r_2}{2\nu_2}\right)\Gamma\left(\frac{r_1}{2\nu_1}\right)},$$

and similarly,

$$\varphi_2 = -\nu_2\left(\frac{d_1}{d_2}\right)^{1/2}\frac{\Gamma\left(\frac{r_1+1}{2\nu_1}\right)\Gamma\left(\frac{r_2+2\nu_2-1}{2\nu_2}\right)}{\Gamma\left(\frac{r_1}{2\nu_1}\right)\Gamma\left(\frac{r_2}{2\nu_2}\right)}.$$

The moment conditions in Lemma 3.2.2 are easily shown to be satisfied for $\nu_k > -(r_k-2)/4$, so putting these values of φ_1 and φ_2 in (3.3) gives the result. □

Note that the condition that $\nu_k > -(r_k-2)/4$ is really a restriction only when $r_k = 1$, in which case we need $\nu_k > 1/4$.

<u>Corollary 3.2.1</u> If $\nu_1 = \nu_2 \equiv \nu$ and $r_1 = r_2 \equiv r$ then

$$\mathrm{ARE}(\hat{Q}_n, -n\log V; \nu; r) = \left[\frac{2\nu\Gamma\left(\frac{r+1}{2\nu}\right)\Gamma\left(\frac{r+2\nu-1}{2\nu}\right)}{r\Gamma^2\left(\frac{r}{2\nu}\right)}\right]^2. \tag{3.5}$$

It is interesting to compare this expression with the Pitman ARE of Randles' interdirection sign statistic, V_n, and Hotelling's T^2 from the multivariate location problem. It turns out that

$$\mathrm{ARE}(\hat{Q}_n, -n\log V; \nu; r) = \frac{\Gamma^2\left(\frac{r+1}{2\nu}\right)}{\Gamma\left(\frac{r}{2\nu}\right)\Gamma\left(\frac{r+2}{2\nu}\right)}\mathrm{ARE}(V_n, T^2; \nu; r).$$

Inserting specific values for r and ν in (3.5), we have that

$$\mathrm{ARE}(\hat{Q}_n, -n\log V; \nu; 1) = \left[\frac{2\nu\Gamma\left(\frac{1}{\nu}\right)}{\Gamma^2\left(\frac{1}{2\nu}\right)}\right]^2,$$

$$\mathrm{ARE}(\hat{Q}_n, -n\log V; 1; r) = \left[\frac{2\Gamma^2\left(\frac{r+1}{2}\right)}{r\Gamma^2\left(\frac{r}{2}\right)}\right]^2,$$

and

$$\mathrm{ARE}(\hat{Q}_n, -n\log V; 0.5; r) = 1,$$

where again it is of interest to note that $\mathrm{ARE}(\hat{Q}_n, -n\log V; 1; r) = \mathrm{ARE}(V_n, T^2; 1; r)^2$. Using the facts that $\lim_{\nu\to\infty}\Gamma(a/\nu)/\Gamma(b/\nu) = b/a$ and $\lim_{\nu\to\infty}\nu/\Gamma(b/\nu) = b$, we also have that

$$\mathrm{ARE}(\hat{Q}_n, -n\log V; \infty; r) = \left[\frac{r}{r+1}\right]^2.$$

Notice that for $r = 1$ (the bivariate case),

$$\mathrm{ARE}(\hat{Q}_n, -n\log V; 0.5; 1) = 1,$$

$$\mathrm{ARE}(\hat{Q}_n, -n\log V; 1; 1) = \frac{4}{\pi^2},$$

and

$$\mathrm{ARE}(\hat{Q}_n, -n\log V; \infty; 1) = \frac{1}{4},$$

which all agree with the values in Table 1.1 and Table 1.2 ($\nu = 0.5$ corresponds to the Laplace, $\nu = 1$ to the normal, and $\nu = \infty$ to the uniform distribution).

As indicated earlier, what follows are several graphs depicting the nature of the Pitman ARE given in Theorem 3.2.4 for various values of $\nu \equiv \nu_1 = \nu_2$. Clearly, $\hat{Q}_n$

does very well when $\nu = 0.1$ (Figure 3.2). Of course, although we can compute and graph the Pitman ARE at the values $r_1 = 1$ and $r_2 = 1$, the expression for the ARE is not valid, so the graph should be ignored in those areas. $\hat{Q}_n$ and $-n \log V$ perform almost equivalently when $\nu = 0.5$ (Figure 3.3), and $-n \log V$ beats $\hat{Q}_n$ when $\nu = 1$ and $\nu = 10$ (Figures 3.4 and 3.5).

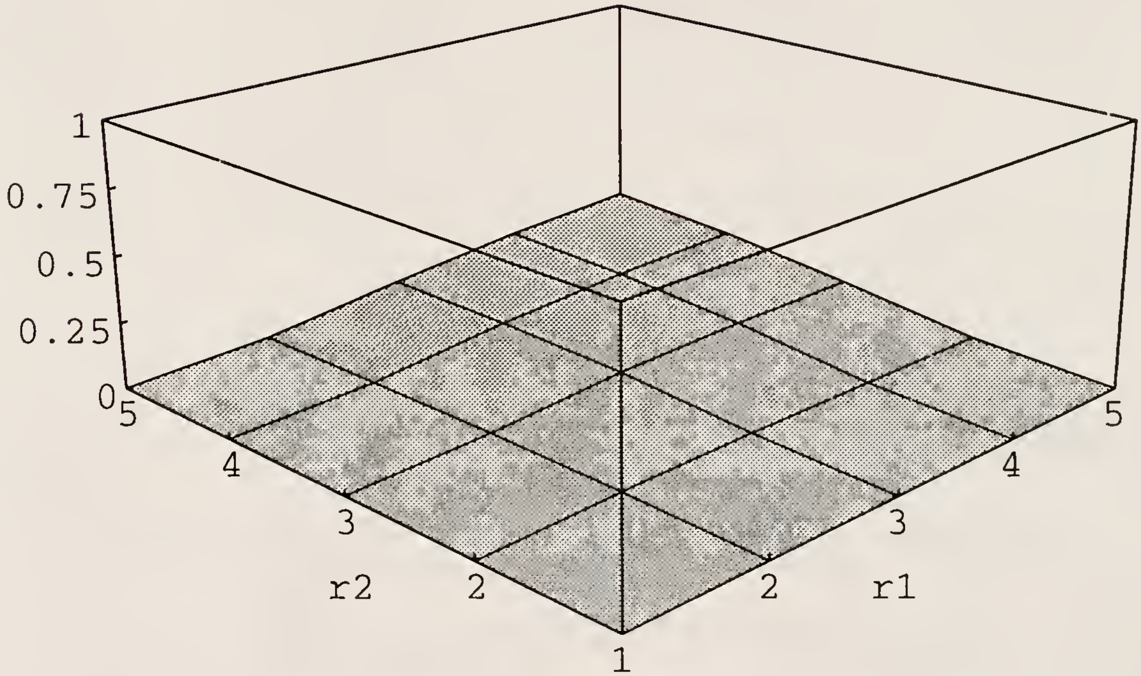

Figure 3.2. $1/(1+\text{ARE}(-n \log V, \hat{Q}_n; \nu = 0.1))$

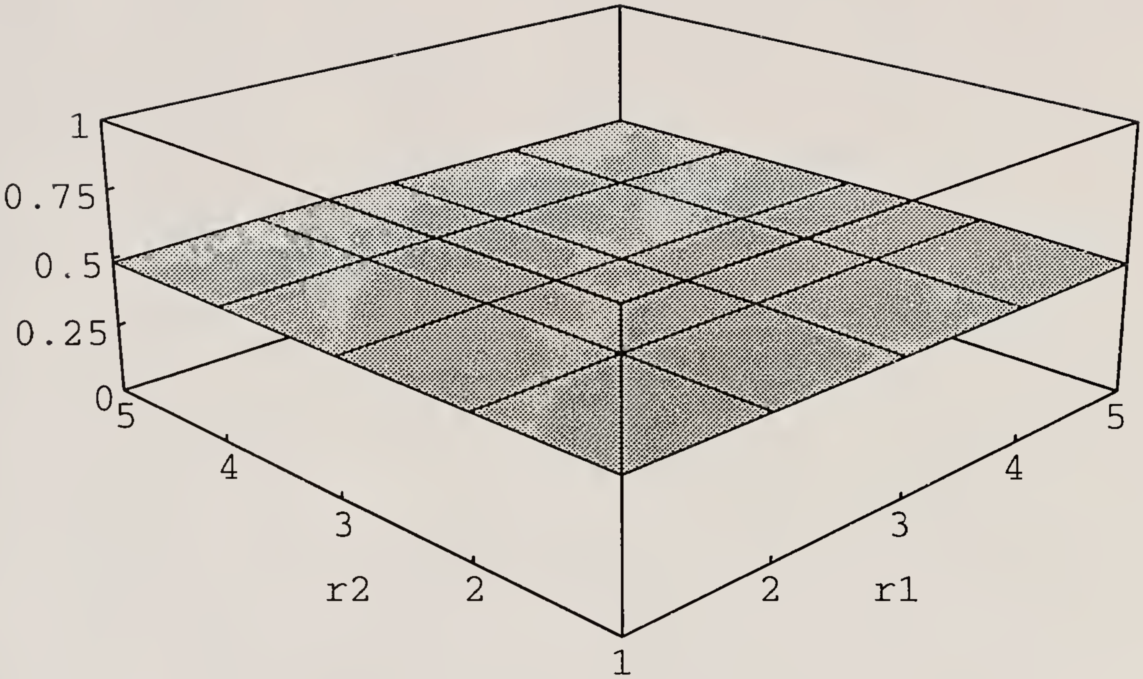

Figure 3.3. $1/(1+\text{ARE}(-n \log V,\, \hat{Q}_n;\, \nu = 0.5))$

Multivariate t-distribution family

The multivariate t-distribution indexed by its degrees of freedom df, is another convenient elliptically symmetric class of distributions. Again we can evaluate the expression for $\text{ARE}(\hat{Q}_n, -n \log V)$ when the underlying distributions of $\boldsymbol{X}^{(1)}$ and $\boldsymbol{X}^{(2)}$

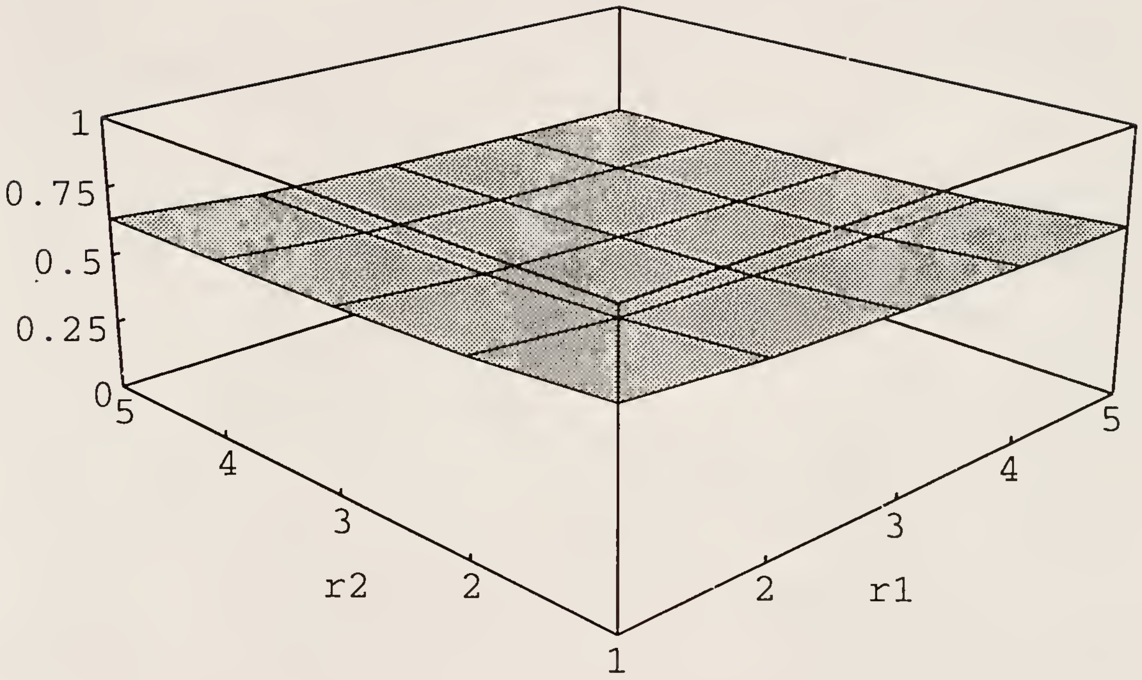

Figure 3.4. $1/(1+\mathrm{ARE}(\hat{Q}_n,\ -n\log V;\ \nu=1))$

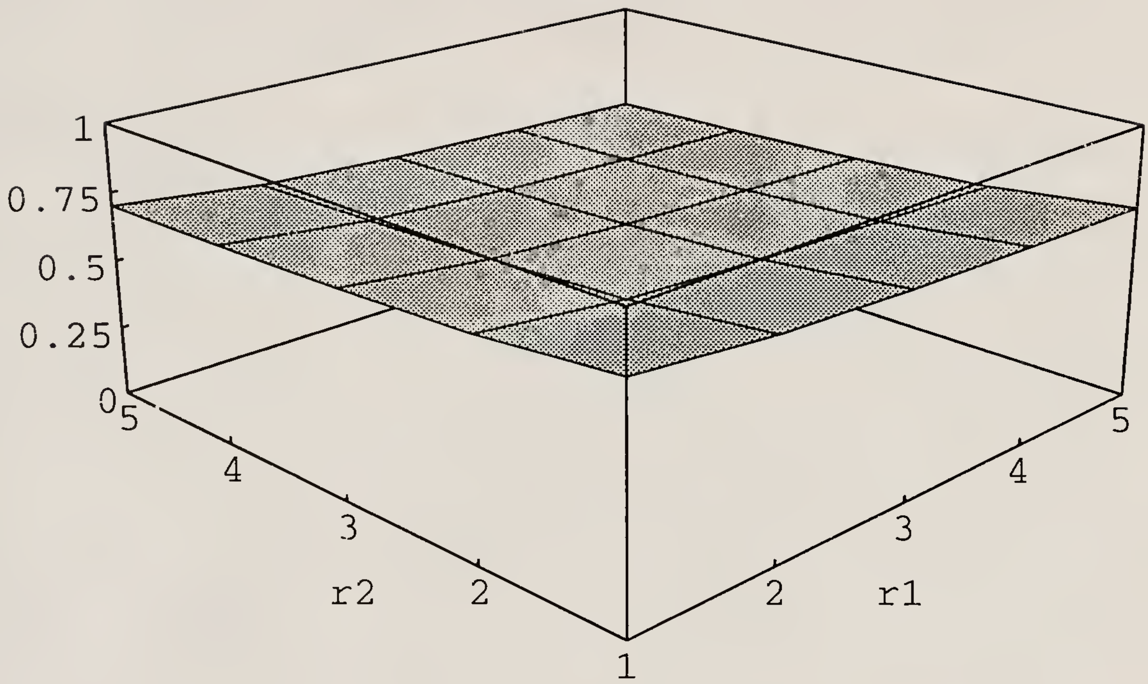

Figure 3.5. $1/(1+\mathrm{ARE}(\hat{Q}_n, -n\log V;\ \nu = 10))$

have varying tail weights. When $df = 1$, this corresponds to the multivariate Cauchy distribution and as $df \to \infty$, the distribution approaches multivariate normality.

Let

$$g_k(t) = \left(1 + \frac{t}{df_k}\right)^{-(df_k + r_k)/2},$$

and

$$C_k = \frac{\Gamma\left(\dfrac{df_k + r_k}{2}\right)}{(\pi df_k)^{r_k/2}\Gamma\left(\dfrac{df_k}{2}\right)}.$$

Then $\boldsymbol{X}^{(k)}$ has a multivariate t-distribution located at the origin with dispersion parameter I_{r_k} and degrees of freedom df_k ($\boldsymbol{X}^{(k)} \sim t(df_k)$). The density function of $(R^{(k)})^2$ is given by:

$$h_k(t) = \frac{1}{df_k B\left(\dfrac{r_k}{2}, \dfrac{df_k}{2}\right)} \left(\frac{t}{df_k}\right)^{r_k/2-1} \left(1 + \frac{t}{df_k}\right)^{-(df_k + r_k)/2}, \tag{3.6}$$

where $B(\cdot, \cdot)$ represents the beta function. Thus $(R^{(k)})^2$ has the same distribution as $df_k(1 - U)/U$ where U is distributed as beta with parameters $r_k/2$ and $df_k/2$. Before calculating the Pitman ARE, we derive a useful moment.

<u>Lemma 3.2.5</u>

$$\mathrm{E}_{H_0}\left[\left((R^{(k)})^2\right)^a \left(1 + \frac{(R^{(k)})^2}{df_k}\right)^b\right] = df_k{}^a \cdot \frac{B\left(\dfrac{r_k + 2a}{2}, \dfrac{df_k - 2b - 2a}{2}\right)}{B\left(\dfrac{r_k}{2}, \dfrac{df_k}{2}\right)},$$

provided $r_k + 2a > 0$ and $df_k > 2(a + b)$.

Proof of Lemma 3.2.5 Let T have the density h given in (3.6). Then

$$\mathrm{E}\left[\left(\frac{T}{df_k}\right)^a\left(1+\frac{T}{df_k}\right)^b\right]$$

$$= \frac{1}{df_k B\left(\frac{r_k}{2}, \frac{df_k}{2}\right)} \int_0^\infty (t/df_k)^{r_k/2-1+a}(1+t/df_k)^{-(df_k+r_k)/2+b}dt$$

$$= \frac{1}{B\left(\frac{r_k}{2}, \frac{df_k}{2}\right)} \int_0^\infty t^{r_k/2-1+a}(1+t)^{-(df_k+r_k)/2+b}dt$$

$$= \frac{1}{B\left(\frac{r_k}{2}, \frac{df_k}{2}\right)} \int_0^\infty t^{(r_k+2a)/2-1}(1+t)^{-((r_k+2a)/2+(df_k-2a-2b)/2)}dt$$

$$= \frac{B\left(\frac{r_k+2a}{2}, \frac{df_k-2a-2b}{2}\right)}{B\left(\frac{r_k}{2}, \frac{df_k}{2}\right)},$$

so long as $r_k + 2a > 0$ and $df_k > 2(a+b)$. □

Theorem 3.2.5 If $\boldsymbol{X}^{(k)} \sim t(df_k)$, with $df_k > 4$, and we assume that $\boldsymbol{M}_1 = \boldsymbol{M}_2'$, then

$$\mathrm{ARE}(\hat{Q}_n, -n\log V; df_1, df_2) = \frac{1}{r_1 r_2}\left(\frac{\Gamma\left(\frac{r_1+1}{2}\right)\Gamma\left(\frac{r_2+1}{2}\right)}{\Gamma\left(\frac{r_1}{2}\right)\Gamma\left(\frac{r_2}{2}\right)}\right)^2$$

$$\times\left(\frac{\Gamma\left(\frac{df_1+1}{2}\right)\Gamma\left(\frac{df_2-1}{2}\right)\left(\frac{df_2}{df_1}\right)^{1/2} + \Gamma\left(\frac{df_1-1}{2}\right)\Gamma\left(\frac{df_2+1}{2}\right)\left(\frac{df_1}{df_2}\right)^{1/2}}{\Gamma\left(\frac{df_1}{2}\right)\Gamma\left(\frac{df_2}{2}\right)}\right)^2.$$

Proof of Theorem 3.2.5 Since

$$\phi_k(t) = g_k'(t)/g_k(t) = -\frac{df_k + r_k}{2df_k}\left(1+\frac{t}{df_k}\right)^{-1},$$

$$\begin{aligned}
\varphi_1 &= \mathrm{E}_{H_0}\left[R^{(2)}\right]\mathrm{E}_{H_0}\left[R^{(1)}\phi_1((R^{(1)})^2)\right] \\
&= -\frac{df_1 + r_1}{2df_1}\mathrm{E}_{H_0}\left[R^{(2)}\right]\mathrm{E}_{H_0}\left[R^{(1)}\left(1 + \frac{(R^{(1)})^2}{df_1}\right)^{-1}\right] \\
&= -\frac{df_1 + r_1}{2df_1}\sqrt{df_2}\frac{B\left(\frac{r_2+1}{2}, \frac{df_2 - 1}{2}\right)}{B\left(\frac{r_2}{2}, \frac{df_2}{2}\right)}\sqrt{df_1}\frac{B\left(\frac{r_1+1}{2}, \frac{df_1+1}{2}\right)}{B\left(\frac{r_1}{2}, \frac{df_1}{2}\right)} \\
&= -\frac{df_1 + r_1}{2}\left(\frac{df_2}{df_1}\right)^{1/2}\frac{B\left(\frac{r_2+1}{2}, \frac{df_2 - 1}{2}\right)B\left(\frac{r_1+1}{2}, \frac{df_1+1}{2}\right)}{B\left(\frac{r_2}{2}, \frac{df_2}{2}\right)B\left(\frac{r_1}{2}, \frac{df_1}{2}\right)} \\
&= \left(\frac{df_2}{df_1}\right)^{1/2}\frac{\Gamma\left(\frac{r_1+1}{2}\right)\Gamma\left(\frac{r_2+1}{2}\right)\Gamma\left(\frac{df_1+1}{2}\right)\Gamma\left(\frac{df_2-1}{2}\right)}{\Gamma\left(\frac{r_1}{2}\right)\Gamma\left(\frac{r_2}{2}\right)\Gamma\left(\frac{df_2}{2}\right)\Gamma\left(\frac{df_1}{2}\right)},
\end{aligned}$$

and similarly,

$$\varphi_2 = \left(\frac{df_1}{df_2}\right)^{1/2}\frac{\Gamma\left(\frac{r_1+1}{2}\right)\Gamma\left(\frac{r_2+1}{2}\right)\Gamma\left(\frac{df_1-1}{2}\right)\Gamma\left(\frac{df_2+1}{2}\right)}{\Gamma\left(\frac{r_1}{2}\right)\Gamma\left(\frac{r_2}{2}\right)\Gamma\left(\frac{df_1}{2}\right)\Gamma\left(\frac{df_2}{2}\right)}.$$

The moment conditions of Lemma 3.2.2 are easily seen to be satisfied if $df_k > 2$, but as the marginal fourth moments must exist for $-n\log V$ to have a limiting distribution, we need $df_k > 4$. Putting these expressions for φ_1 and φ_2 in (3.3) gives the result. □

Corollary 3.2.2 If $df_1 = df_2 \equiv df$ and $r_1 = r_2 \equiv r$ then

$$\mathrm{ARE}(\hat{Q}_n, -n\log V; df; r) = \left[\frac{2\Gamma^2\left(\frac{r+1}{2}\right)\Gamma\left(\frac{df+1}{2}\right)\Gamma\left(\frac{df-1}{2}\right)}{r\Gamma^2\left(\frac{r}{2}\right)\Gamma^2\left(\frac{df}{2}\right)}\right]^2. \tag{3.7}$$

Interestingly, the expression (3.7) can be factored into the Pitman ARE under multivariate normality times a quantity involving only df, which (obviously) goes to one

as $df \to \infty$. Thus, for example, putting $df = 5$ yields

$$\mathrm{ARE}(\hat{Q}_n, -n\log V; 5; r) = \frac{1024}{81\pi^2}\left[\frac{2\Gamma^2\left(\dfrac{r+1}{2}\right)}{r\Gamma^2\left(\dfrac{r}{2}\right)}\right]^2,$$

where $1024/81\pi^2 = 1.2809$ or putting $r = 1$ yields

$$\mathrm{ARE}(\hat{Q}_n, -n\log V; df; 1) = \frac{4}{\pi^2}\left[\frac{\Gamma\left(\dfrac{df+1}{2}\right)\Gamma\left(\dfrac{df-1}{2}\right)}{\Gamma^2\left(\dfrac{df}{2}\right)}\right]^2.$$

Again we include some graphs of the Pitman ARE given in Theorem 3.2.5 for various values of $df \equiv df_1 = df_2$. We could calculate the Pitman ARE when $df < 5$, but since $-n\log V$ is guaranteed to have a limiting distribution only when $df \geq 5$, the value would be meaningless. Hence we consider only $df \geq 5$. Of course, we anticipate that $\hat{Q}_n$ is vastly superior to $-n\log V$ when $df < 5$, but we do not have a way to quantify their relative performance in this instance. (We do include $df = 1$ in the simulation study, which will demonstrate if our intuitive feeling is borne out.) When $df = 5$ (Figure 3.6), we see that $\hat{Q}_n$ and $-n\log V$ perform essentially the same. When $df = 10$ and $df = 100$ (Figures 3.7 and 3.8), $-n\log V$ beats $\hat{Q}_n$. In fact, for $df = 10$, the Pitman ARE is already only 1.057 times that of the multivariate normal Pitman ARE, for which $-n\log V$ is optimal.

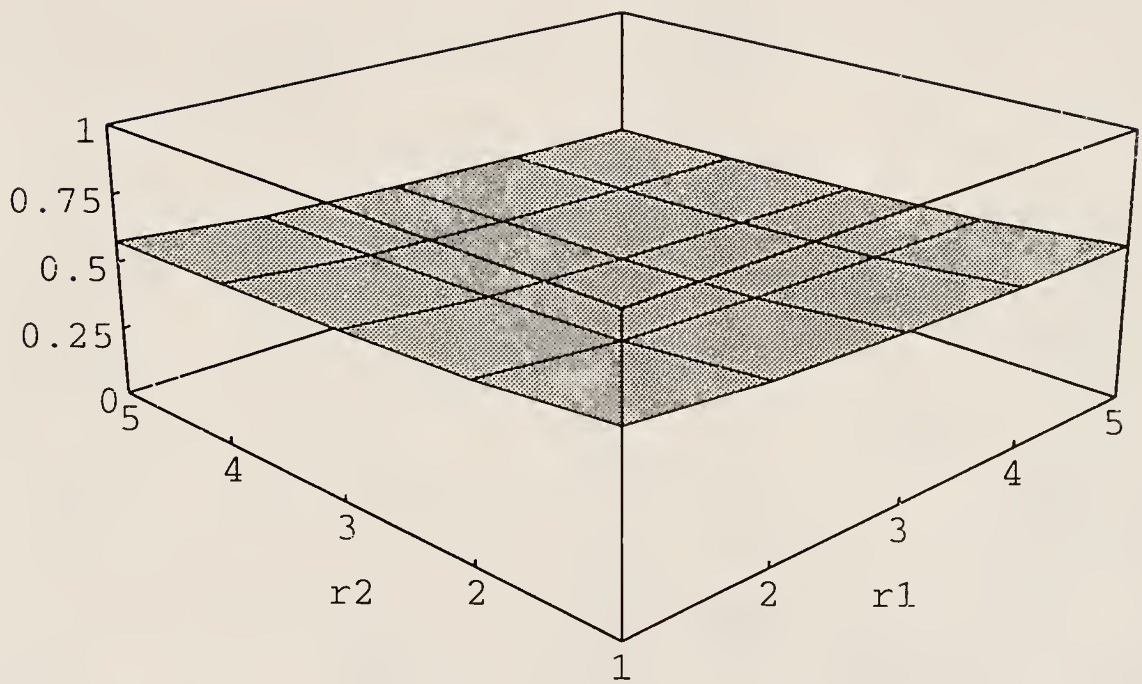

Figure 3.6. $1/(1+\mathrm{ARE}(-n \log V,\, \hat{Q}_n;\, df = 5))$

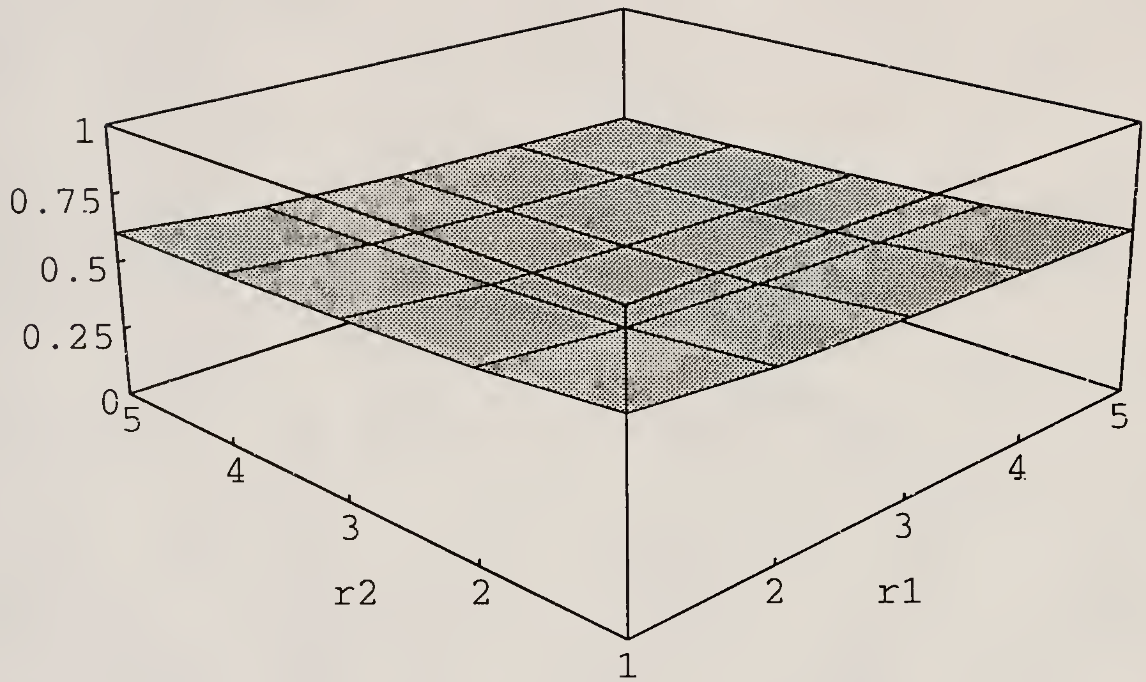

Figure 3.7. $1/(1+\mathrm{ARE}(\hat{Q}_n, -n\log V; df = 10))$

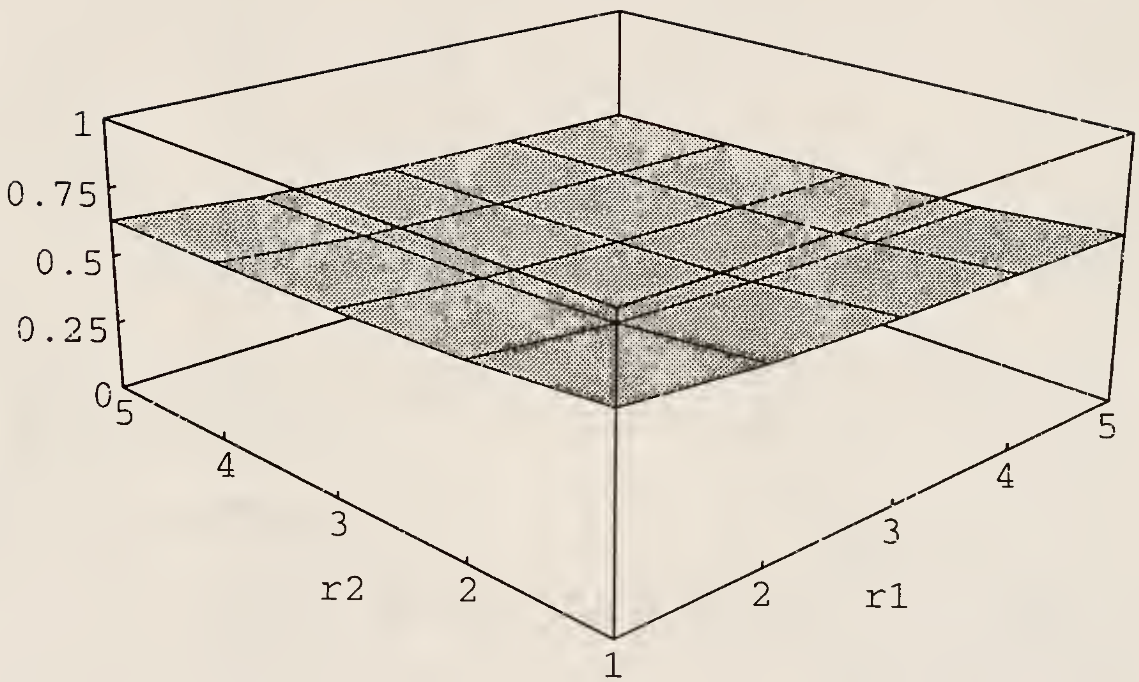

Figure 3.8. $1/(1+\mathrm{ARE}(\hat{Q}_n,\ -n\log V;\ df = 100))$

3.3 Model 2

A model proposed by Puri and Sen (1971) is given by

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}^{(1)} \\ \boldsymbol{X}^{(2)} \end{pmatrix} = \begin{pmatrix} \boldsymbol{Y}^{(1)} + \Delta \boldsymbol{Z}^{(1)} \\ \boldsymbol{Y}^{(2)} + \Delta \boldsymbol{Z}^{(2)} \end{pmatrix}$$

$$= \begin{pmatrix} \boldsymbol{Y}^{(1)} \\ \boldsymbol{Y}^{(2)} \end{pmatrix} + \Delta \begin{pmatrix} \boldsymbol{Z}^{(1)} \\ \boldsymbol{Z}^{(2)} \end{pmatrix}$$

$$= \boldsymbol{Y} + \Delta \boldsymbol{Z},$$

where $\boldsymbol{Y}^{(1)}$, $\boldsymbol{Y}^{(2)}$ and $\boldsymbol{Z}$ are mutually independent, and the matrix $\mathrm{Cov}_{H_0}\left[\boldsymbol{Z}^{(1)}, \boldsymbol{Z}^{(2)}\right]$ consists entirely of nonzero elements (although it appears that it is sufficient for it not to be the zero matrix). They state that such a model may prove useful in analyzing group tests in psychology. For example, the outcomes of two reading tests and two math tests can be described by a (linear) combination of individual group factors pertaining to the reading or mathematical abilities and common factors corresponding to intelligence or comprehension. In general, the distribution of $\boldsymbol{X}$ is determined by using a convolution formula since it is a sum of two independent random vectors. However, obtaining a closed form expression for the density function, an integral step in being able to work out the details related to contiguity, is typically non-trivial. An exception is when $\boldsymbol{Y}^{(1)}$, $\boldsymbol{Y}^{(2)}$ and $\boldsymbol{Z}$ have multivariate normal distributions. Thus, if $\boldsymbol{Y}^{(1)} \sim \mathrm{MVN}(\boldsymbol{0}, \boldsymbol{I}_{r_1})$, $\boldsymbol{Y}^{(2)} \sim \mathrm{MVN}(\boldsymbol{0}, \boldsymbol{I}_{r_2})$, and $\boldsymbol{Z} \sim \mathrm{MVN}(\boldsymbol{0}, \Sigma)$, then $\boldsymbol{X} \sim \mathrm{MVN}(\boldsymbol{0}, \boldsymbol{I}_{r_1+r_2} + \Delta^2\Sigma)$.

Since the determination of contiguity and computation of the limiting distributions of $-n \log V$, $\hat{Q}_n$, and $-n \log S^{J_0}$ under this sequence of multivariate normal alternatives is virtually identical to the computation under alternatives described by

Model 1, the details will be omitted. The end result is that

$$-n\log V \xrightarrow{d} \chi^2_{r_1r_2}\left(\Delta_0^2\mathrm{vec}\,(\boldsymbol{\Sigma}_{12})'\,\mathrm{vec}\,(\boldsymbol{\Sigma}_{12})\right),$$

$$-n\log S^{J_0} \xrightarrow{d} \chi^2_{r_1r_2}\left(\frac{4}{\pi^2}\Delta_0^2\mathrm{vec}\,(\boldsymbol{\Sigma}_{12})'\,\mathrm{vec}\,(\boldsymbol{\Sigma}_{12})\right),$$

and

$$\hat{Q}_n \xrightarrow{d} \chi^2_{r_1r_2}\left(\left[\frac{2\Gamma\left(\frac{r_1+1}{2}\right)\Gamma\left(\frac{r_2+1}{2}\right)}{r_1r_2\Gamma\left(\frac{r_1}{2}\right)\Gamma\left(\frac{r_2}{2}\right)}\right]^2\Delta_0^2\mathrm{vec}\,(\boldsymbol{\Sigma}_{12})'\,\mathrm{vec}\,(\boldsymbol{\Sigma}_{12})\right).$$

From these expressions, we observe that the Pitman ARE's will be identical to those derived from Model 1 when the underlying distribution is multivariate normal. This is further indication that the actual form of the model used is somewhat irrelevant when dealing with local alternatives.

CHAPTER 4
MONTE CARLO STUDY

4.1 Methods

All simulation programs were written in the C programming language. Previously written routines acquired from various sources were combined with original source code to complete the main procedure. Outside sources included a large archive of software maintained by AT&T called Netlib `<netlib@research.att.com>` and another archive of statistically related software maintained by Michael Meyer at Carnegie Mellon University called Statlib `<statlib@lib.stat.cmu.edu>`. Included in the non-original code used are parts of the following libraries: `c/meschach`—a set of functions which do numerical linear algebra, dense and sparse, with permutations, error handling and input/output by David E. Stewart `<des@thrain.anu.edu.au>`, `c/cephes`—a set of special math functions and IEEE floating point arithmetic by Stephen L. Moshier `<moshier@world.std.com>` and `ranlib-c`—a set of random variate generators translated from FORTRAN by Barry Brown `<bwb@odin.mda.uth.tmc.edu>`. Also used are `chisq.c`, `f.c` and `z.c`, which are functions written by Gary Perlman to compute probabilities and percentiles of the chi-square, F and normal distributions and `L1.f`, a FORTRAN routine based on an algorithm by Barrodale and Roberts (1974) to compute the least absolute value solution to an over-determined system of equations (personally translated to C). The final program was compiled using `gcc` (GNU project C compiler v2.4) on a SPARC 10. Several programs which are of interest are included in Appendix C.

The two distribution types used were the exponential class and multivariate t described earlier. The method for generating observations from these distributions has three parts (see Johnson, 1987). First, a vector of iid $N(0,1)$ random variables is generated. Second, the vector is divided by its Euclidean norm, which results in a vector uniformly distributed on the unit hypersphere. Third, multiplication by a positive random scalar with the appropriate distribution ($\sqrt{d_k}\,\mathrm{Gamma}(r/(2\nu_k),1)^{1/(2\nu_k)}$ for the exponential class and $\sqrt{df_k/\chi^2_{df_k}}$ for the multivariate t) yields the desired multivariate observation. Model 1 (3.1) was then used to generate the dependence structure. For ease of comparison, we restricted the study to cases where $r_1 = r_2 \equiv r$ and the underlying distribution types were identical for each set of variables. Specifically, for the dimensions $r = 1,2$, we used the distributions $\nu = 0.1, 0.5, 1, 10$ in the exponential class and $df = 1,5$ in the multivariate t. For $r = 3$, we considered only the exponential class with $\nu = 0.1, 0.5, 1$. The sample size, n, was kept at 30 and the number of repetitions at each setting was 2500 when $r = 1,2$. Because of the increased computing time needed when $r = 3$, the number of repetitions was decreased to 1000. To gauge the precision of the empirical powers computed, we provide a table summarizing the maximum estimated standard errors of the empirical power over various levels of the true power.

4.2 Statistics Compared

Of course $-n\log V$, $-n\log S^{J_0}$, and $\hat{Q}_n$ were included in the study, but numerous other normal theory tests not explicitly investigated in this thesis were also added to judge $\hat{Q}_n$ against. In particular, the standard multivariate tests used in the SAS^{tm} procedure PROC CANCORR and their variants were considered. These tests are all based on the sample canonical correlations $(\hat{\rho}_1,\ldots,\hat{\rho}_{r_1})$, where $\hat{\rho}_1^2 > \cdots > \hat{\rho}_{r_1}^2$

Table 4.1. Maximum Estimated Standard Errors for Empirical Power

	Repetitions	
Power	1000	2500
0.00 – 0.05	0.0069	0 0044
0.05 – 0.10	0.0095	0.0060
0.10 – 0.25	0.0137	0.0087
0.25 – 0.75	0.0158	0.0100
0.75 – 0.90	0.0137	0.0087
0.90 – 0.95	0.0095	0.0060
0.95 – 1.00	0.0069	0.0044

are the eigenvalues of $\boldsymbol{S} \equiv \boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{12}'$. Several of the tests can be approximated both by an F distribution and a chi-square distribution. We present both forms for comparison. First is $-n\log V$, which multiplied by the Bartlett correction factor $1-(r_1+r_2+3)/(2n)$ (Box, 1949) is labeled as LX. The F approximation (a different transformation of V) is naturally labeled LF. VF and VX are F and chi-square approximations, respectively, for Pillai's trace $\sum_{i=1}^{r_1}\hat{\rho}_i^2$. Likewise UF and UX are F and chi-square approximations, respectively, for the Lawley-Hotelling trace $\sum_{i=1}^{r_1}\hat{\rho}_i^2/(1-\hat{\rho}_i^2)$. The last normal theory test is RF, an upper bound on an F approximation to Roy's greatest root $\hat{\rho}_1^2/(1-\hat{\rho}_1^2)$. For consistency, we label $-n\log S^{J_0}$ as PS.

The Oja median (Oja, 1983) was used in $\hat{Q}_n$, which we will henceforth call Q1, to estimate the nuisance parameters $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$. This generalized median, which is the point minimizing the sum of the volumes of all p-dimensional simplexes formed from $p-1$ sample points and itself, is equivariant under the group $\mathcal{G}$ and asymptotically normal with rate $\sqrt{n}$ (Oja & Niinimaa, 1985). When $p=1$, the Oja median is just

the usual univariate median. The method used to compute it is based on the L_1-norm formulation of the minimization as given in Niinimaa (1992).

The final statistic, Q2, is defined as

$$\mathrm{Q2} = \frac{r_1 r_2}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \hat{\boldsymbol{U}}_i^{(1)\prime} \hat{\boldsymbol{U}}_j^{(1)} \cdot \hat{\boldsymbol{U}}_i^{(2)\prime} \hat{\boldsymbol{U}}_j^{(2)},$$

where

$$\hat{\boldsymbol{U}}^{(1)} = \frac{\hat{\Sigma}_{11}^{-1/2}(\boldsymbol{X}^{(1)} - \hat{\boldsymbol{\theta}}_1)}{\sqrt{(\boldsymbol{X}^{(1)} - \hat{\boldsymbol{\theta}}_1)'\hat{\Sigma}_{11}^{-1}(\boldsymbol{X}^{(1)} - \hat{\boldsymbol{\theta}}_1)}}$$

and

$$\hat{\boldsymbol{U}}^{(2)} = \frac{\hat{\Sigma}_{22}^{-1/2}(\boldsymbol{X}^{(2)} - \hat{\boldsymbol{\theta}}_2)}{\sqrt{(\boldsymbol{X}^{(2)} - \hat{\boldsymbol{\theta}}_2)'\hat{\Sigma}_{22}^{-1}(\boldsymbol{X}^{(2)} - \hat{\boldsymbol{\theta}}_2)}},$$

and was included because it is asymptotically equivalent to Q1, but is much simpler computationally. Robust M-estimates of Σ_{11} and Σ_{22} as described by Randles, Broffitt, Ramberg, and Hogg (1978) were used in Q2 as well as the same Oja median statistics used in Q1 to estimate $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$. Thus Q2 is invariant under $\mathcal{G}$ and it is hoped that it will be as robust as Q1.

4.3 Results

The outcome of the Monte Carlo study is presented graphically in a series of figures at the end of the chapter. However, the values used in these figures are in Appendix C in tabular form as well. In order to facilitate understanding of the simulation results, we make some general comments for specific cases.

When $r = 1$ (the bivariate case) the results are not unexpected. Since Q1 and Q2 are essentially equivalent to q' (a very robust statistic), it is easy to understand why, for the heavy-tailed distributions, (see Figures 4.1 and 4.5 for graphs of $\nu = 0.1$ and $df = 1$) both Q1 and Q2 do better than the normal theory tests. It is

important to keep in mind that by saying "better", we mean that, not only does the test have higher power over the alternative, but that the test has at least come close to maintaining the designated nominal level of 0.05. For tests that do not achieve the latter criterion, it is difficult to compare them with competing procedures. In general however, it seems prudent to be biased in favor of procedures which maintain the desired level versus procedures which have higher power but do not maintain the the nominal level. Thus although there are instances where Q1 and Q2 are "beat" by other procedures at various points in the alternative, for the most part Q1 and Q2 are much better at maintaining the designated 0.05 level, indicating their favorability. This is the case here, and in general, for heavy-tailed distributions like the multivariate t with $df = 1$. An interesting observation is that although PS is asymptotically equivalent to q', it doesn't do as well as Q1 and Q2 for these distributions. In fact, it is uniformly worse (conservative) than both Q1 and Q2. One reason may be that the central limit theorem doesn't work quite as fast on the log of a sum (like PS) as it does on a strict sum of terms (like Q1 and Q2). We note that in this case the normal theory tests perform almost identically, with the exception of UX, which is too liberal. Hence for $r = 1$ we do not differentiate among the normal tests, except to exclude UX. As the distributions become lighter tailed (see Figures 4.2, 4.3, 4.4, and 4.6 for graphs of $\nu = 0.5,\ 1,\ 10$ and $df = 5$), the normal tests are clearly better than than Q1, Q2, and PS.

When $r = 2$, we examine the competitors showings in greater detail. Since RF and UX are consistently above the 0.05 level by a large margin, while the rest of the normal theory tests, although they may also exceed it, are less liberal and perform comparably, we will not differentiate among the normal theory tests except to exclude RF and UX. For $\nu = 0.1$ and $df = 1$, (see Figures 4.7 and 4.11), both Q1 and Q2 perform well, maintaining the 0.05 significance level and showing a steep increase in power. PS comes close to doing as well, but overshoots the 0.05 level and has

power slightly below Q1 and Q2. None of the normal theory tests do very well. In fact, for $\nu = 0.1$, they all have uniformly much lower power than Q1 and Q2 while greatly exceeding the 0.05 nominal level. For $\nu = 0.5$ and $df = 5$ (see Figures 4.8 and 4.12), Q1 and Q2 seem to do equally as well as the normal theory tests, with PS again exceeding the 0.05 level and having power somewhat less than the others. For $\nu = 1, 10$ (see Figures 4.9 and 4.10), Q1 and Q2 start off very competitive, but have decreased power relative to the normal tests as the dependency is increased. PS again is uniformly worse than the others, having exceeded the 0.05 level and having lower power over the alternative.

When $r = 3$, the situation remains essentially the same. Referring to Figures 4.13, 4.14, and 4.15, RF and UX are uniformly bad with Q1 and Q2 doing slightly worse than the other normal tests for $\nu = 1$ but competitive for $\nu = 0.5$ and dominant for $\nu = 0.1$.

In general, it appears that Q1 and Q2 do a much better job of maintaining their nominal level than PS or any of the normal theory tests. Q1 and Q2 are consistently very close to maintaining the 0.05 nominal level where the others vary widely either above or below. This implies that the small sample applicability of Q1 or Q2 is very good. We might add that most of the simulation results presented here are basically in agreement with the Pitman ARE's derived in Chapter 3. For example, compare the essentially equivalent performances of Q1 and LX when $\nu = 0.5$ with the associated Pitman ARE (see Figure 3.3).

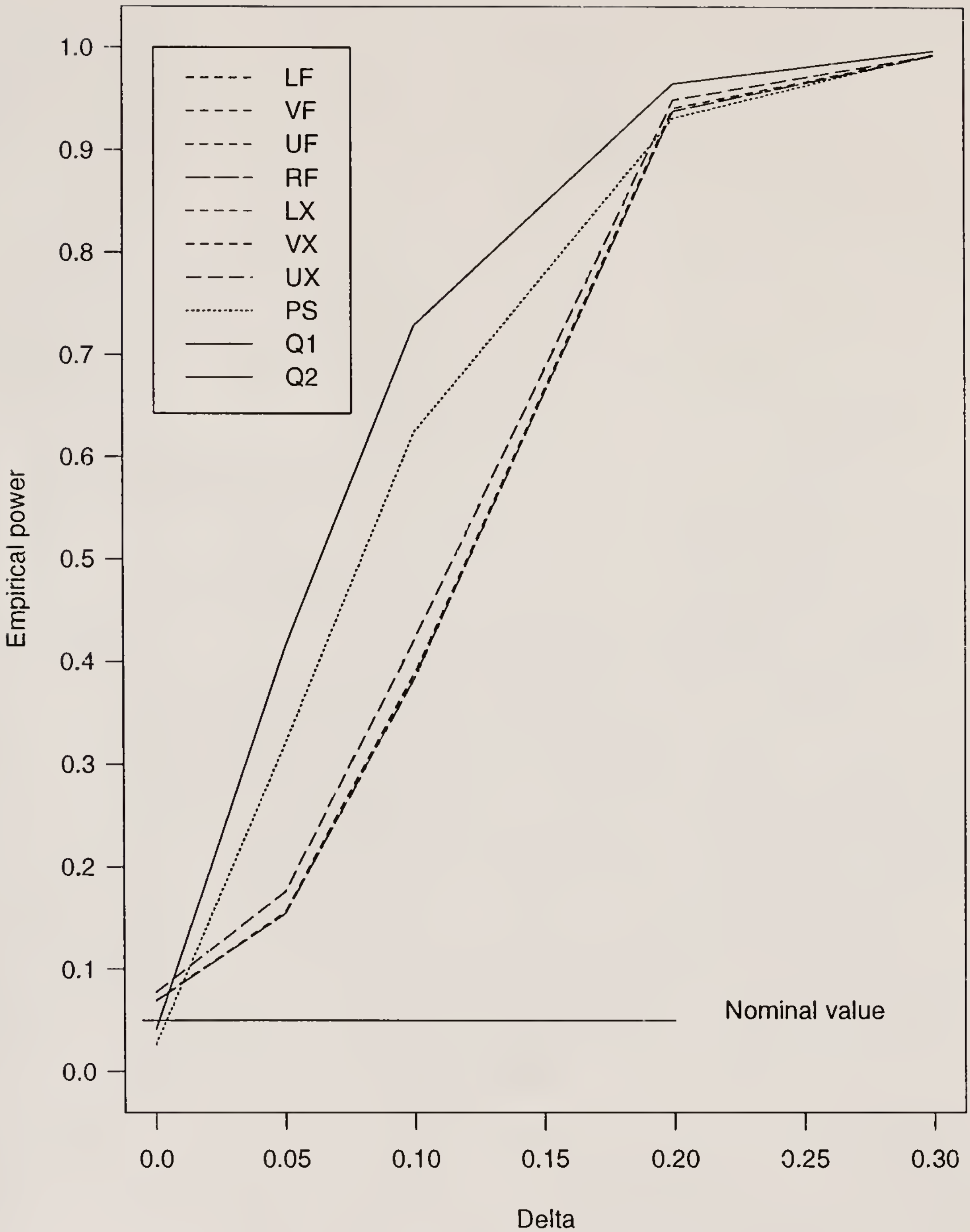

Figure 4.1. $r = 1$, $n = 30$, $\nu = 0.1$, reps $= 2500$

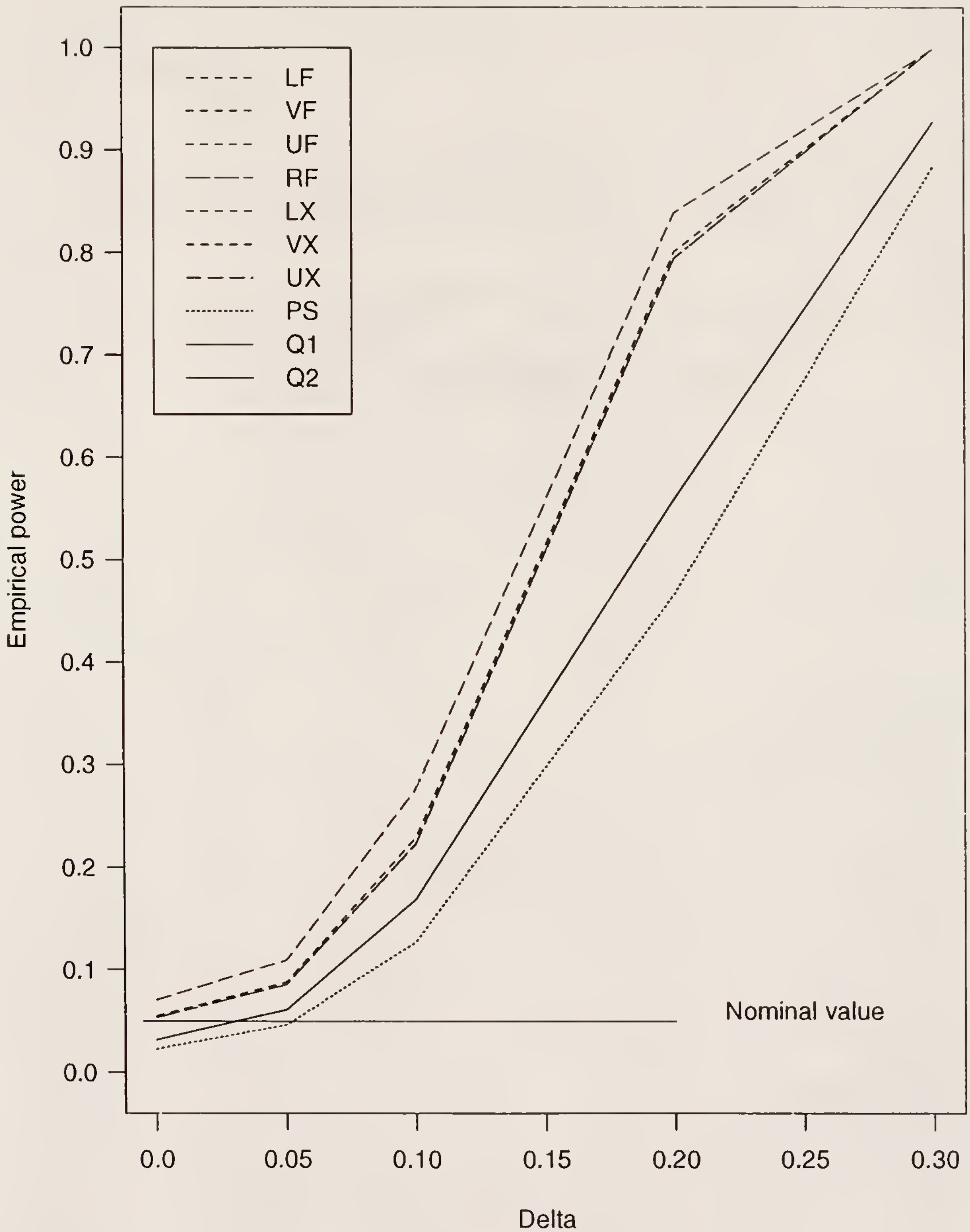

Figure 4.2. $r = 1$, $n = 30$, $\nu = 0.5$, reps $= 2500$

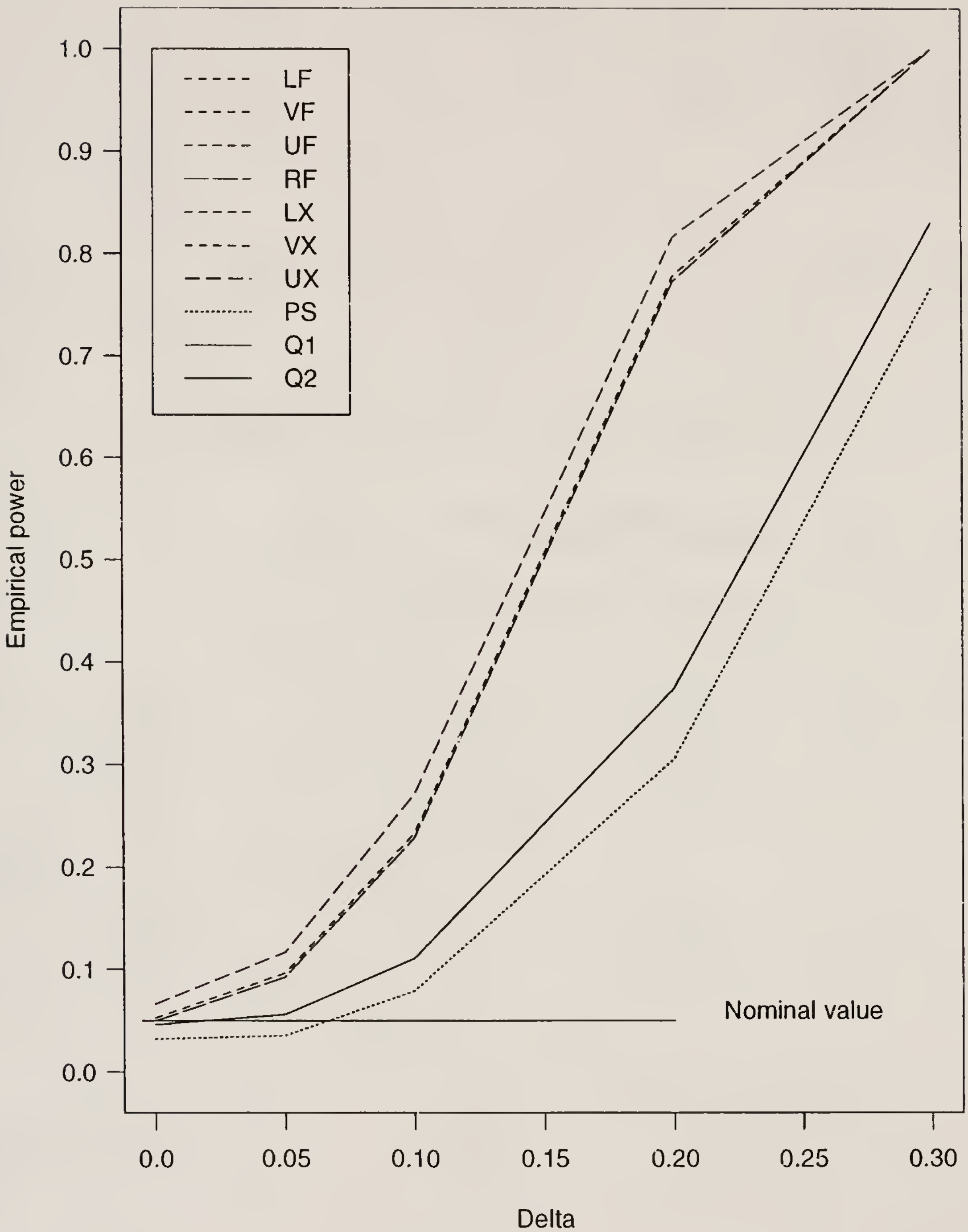

Figure 4.3. $r = 1$, $n = 30$, $\nu = 1$, reps = 2500

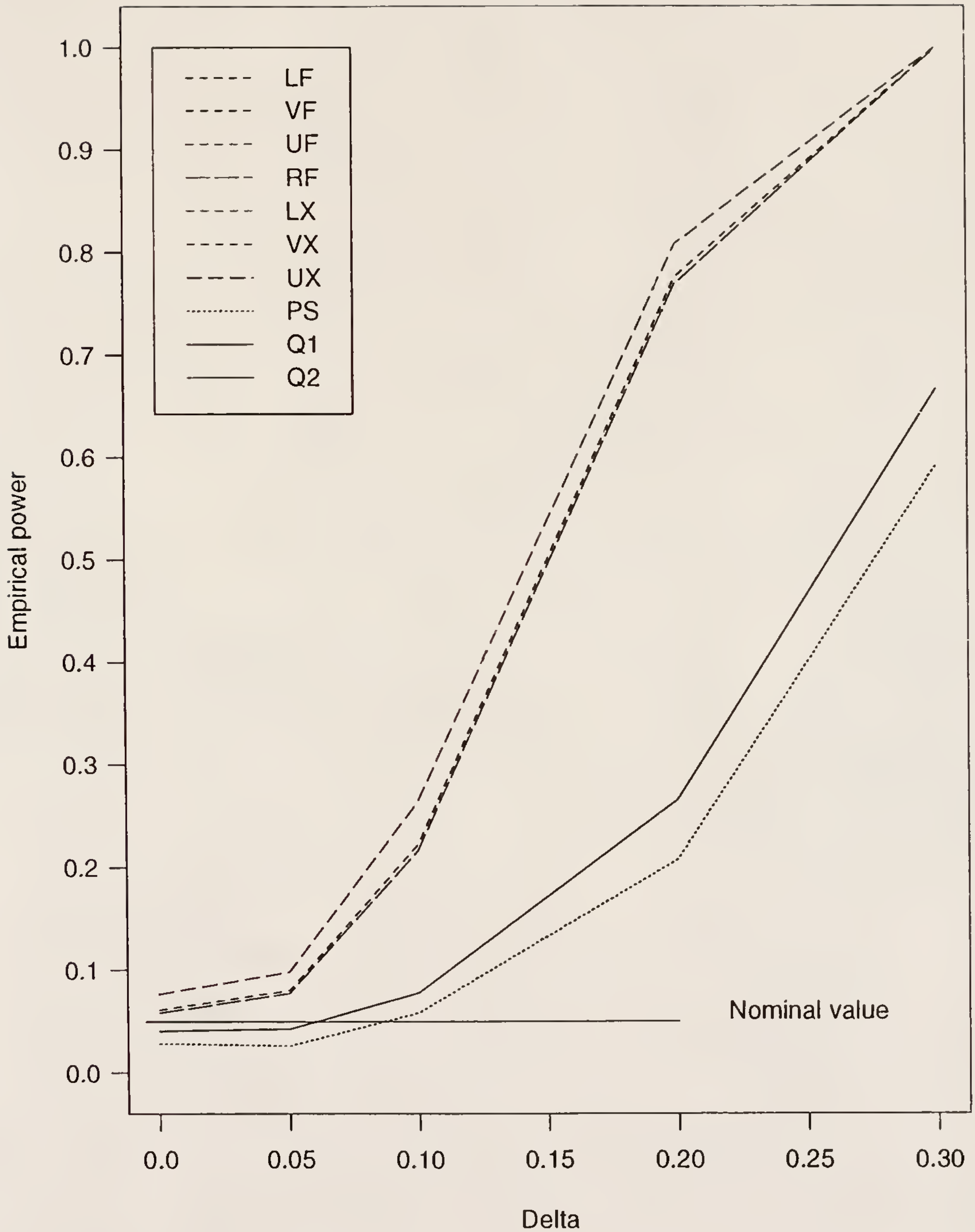

Figure 4.4. $r = 1$, $n = 30$, $\nu = 10$, reps = 2500

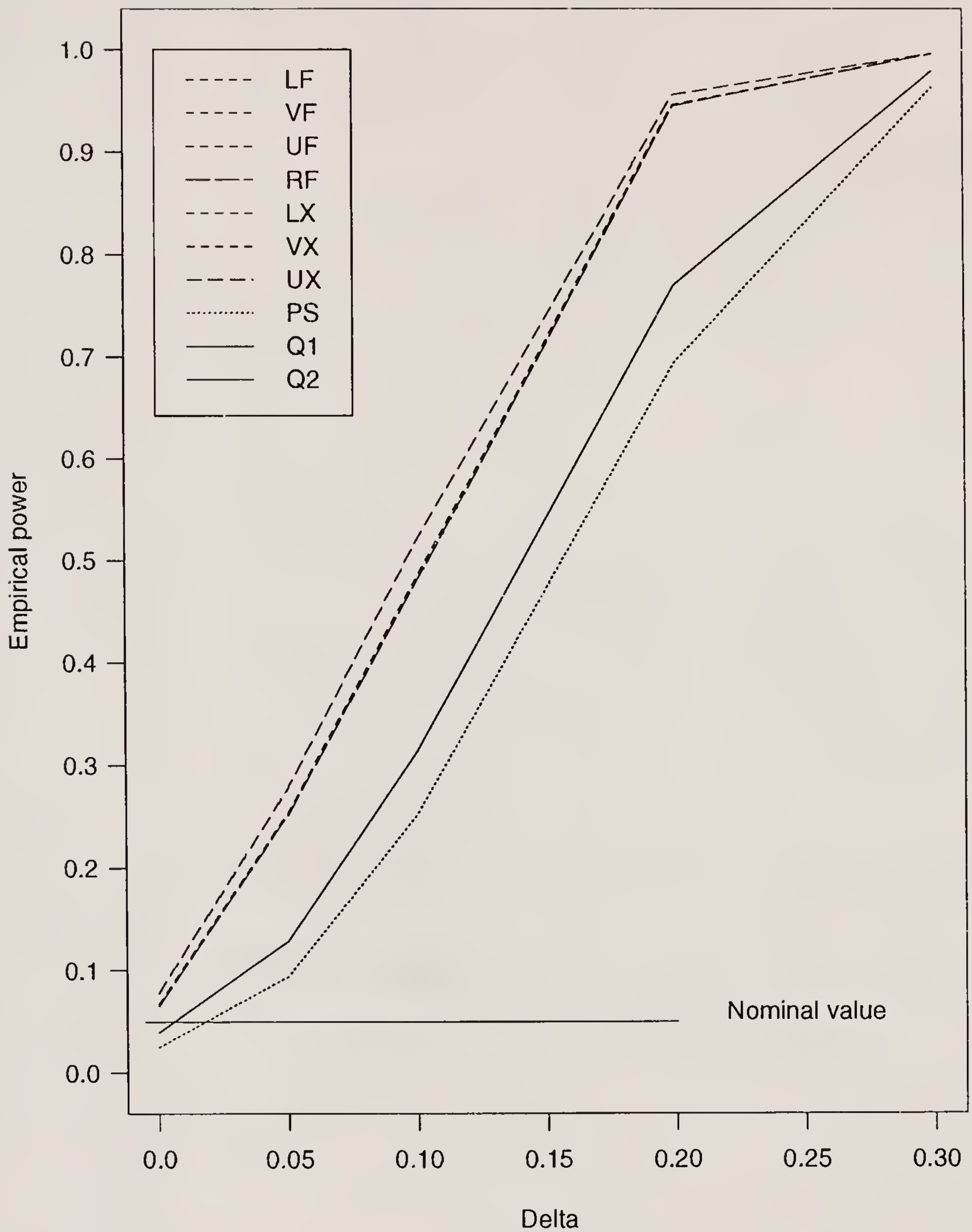

Figure 4.5. $r = 1$, $n = 30$, $df = 1$, reps = 2500

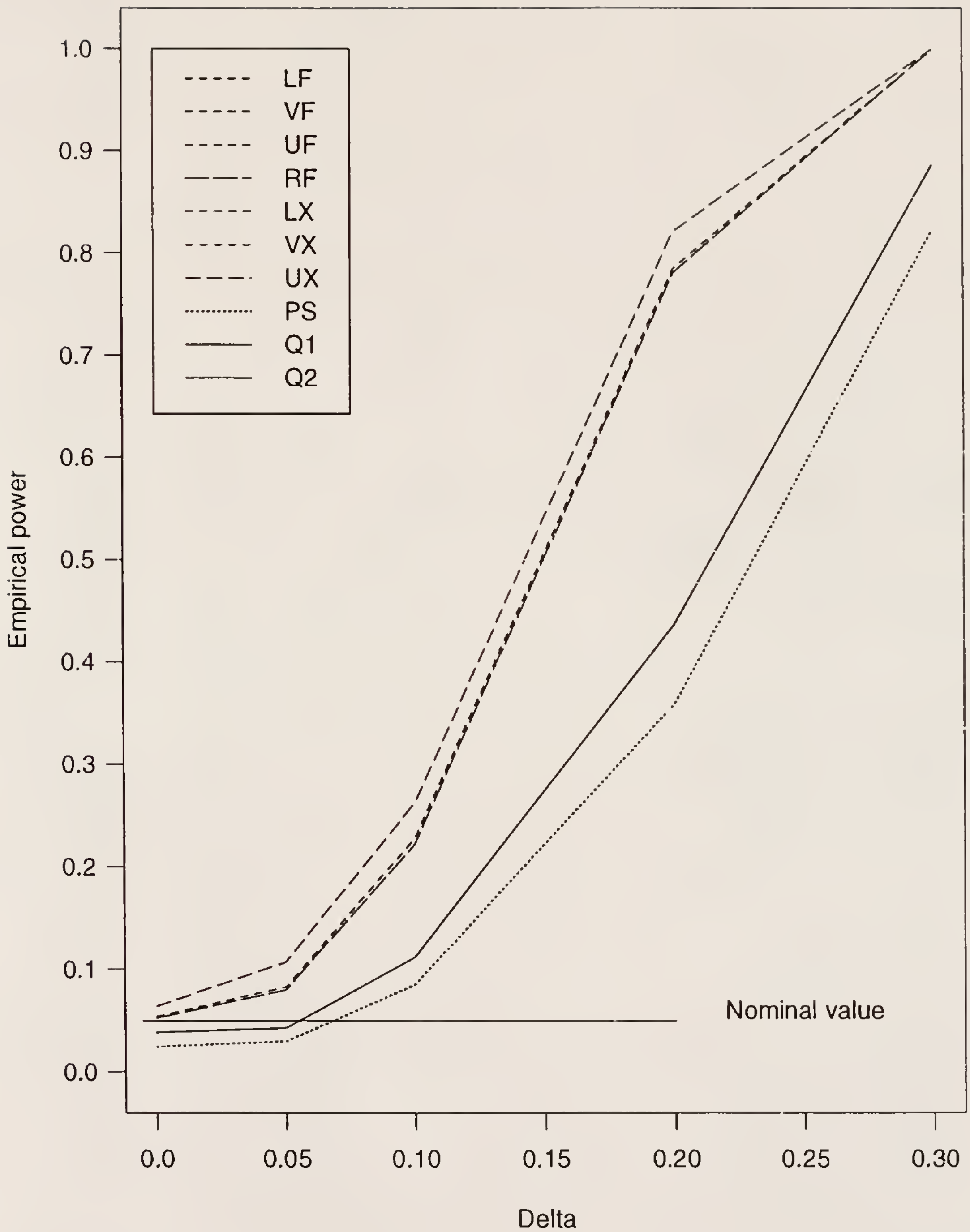

Figure 4.6. $r = 1$, $n = 30$, $df = 5$, reps $= 2500$

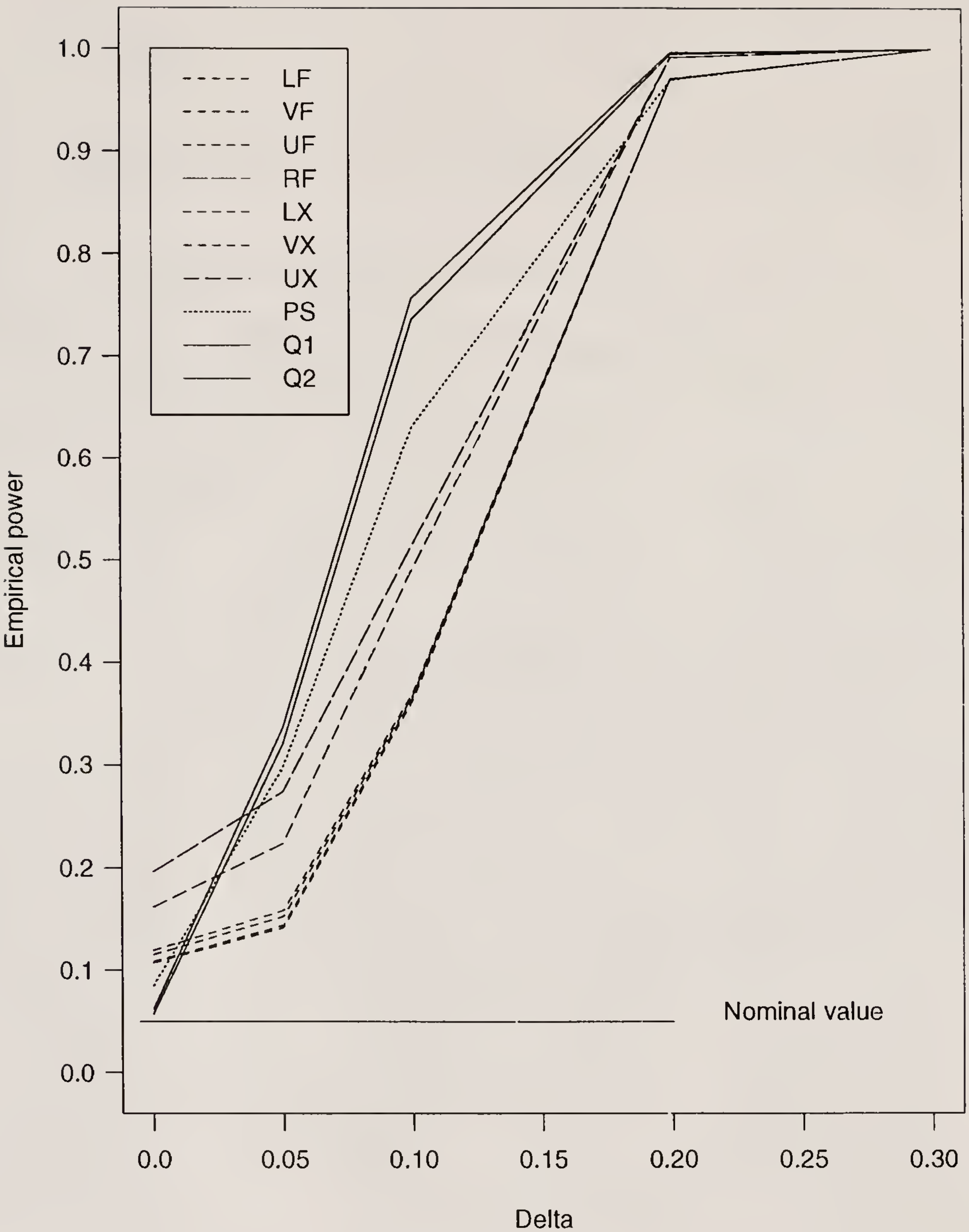

Figure 4.7. $r = 2$, $n = 30$, $\nu = 0.1$, reps = 2500

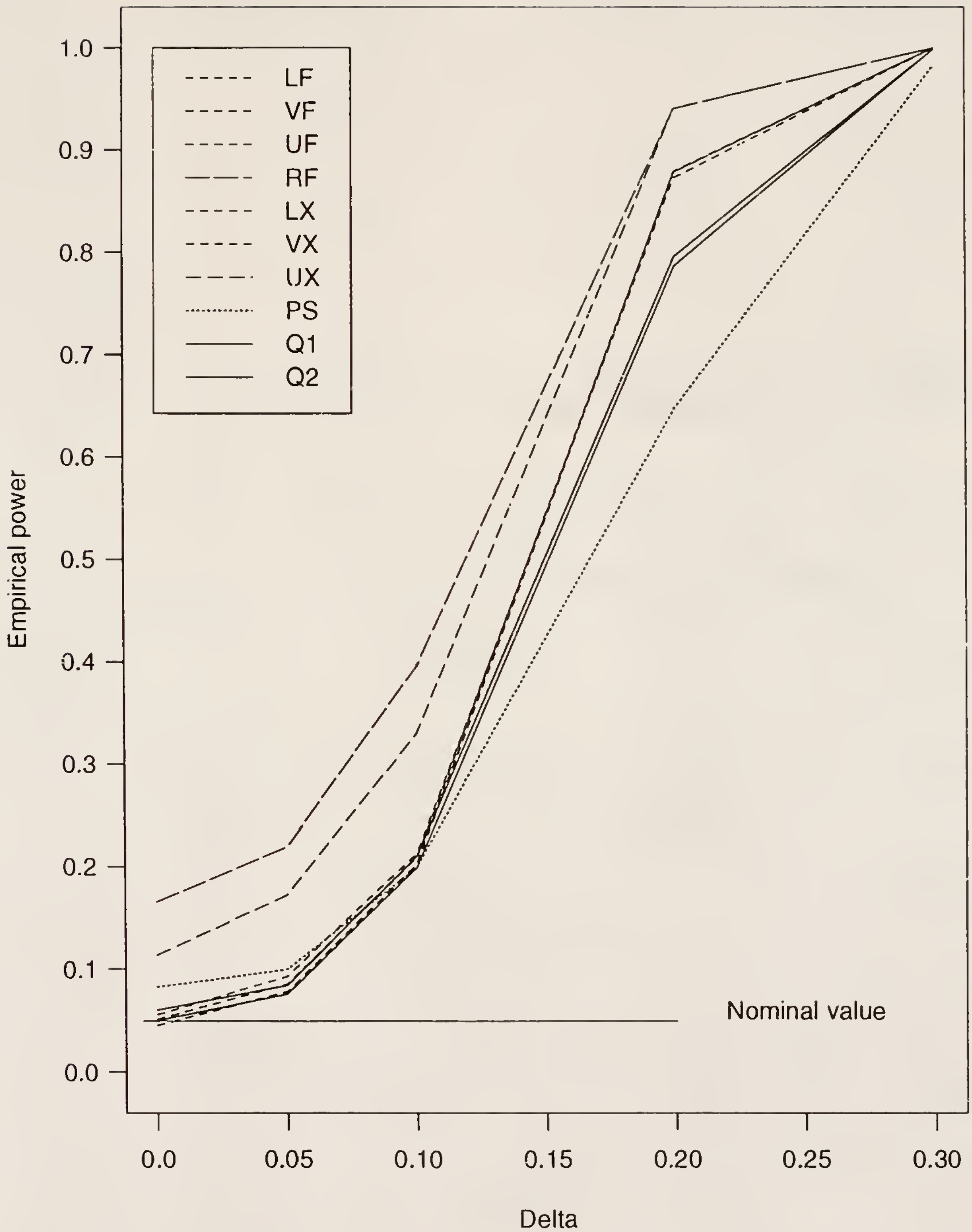

Figure 4.8. $r = 2$, $n = 30$, $\nu = 0.5$, reps = 2500

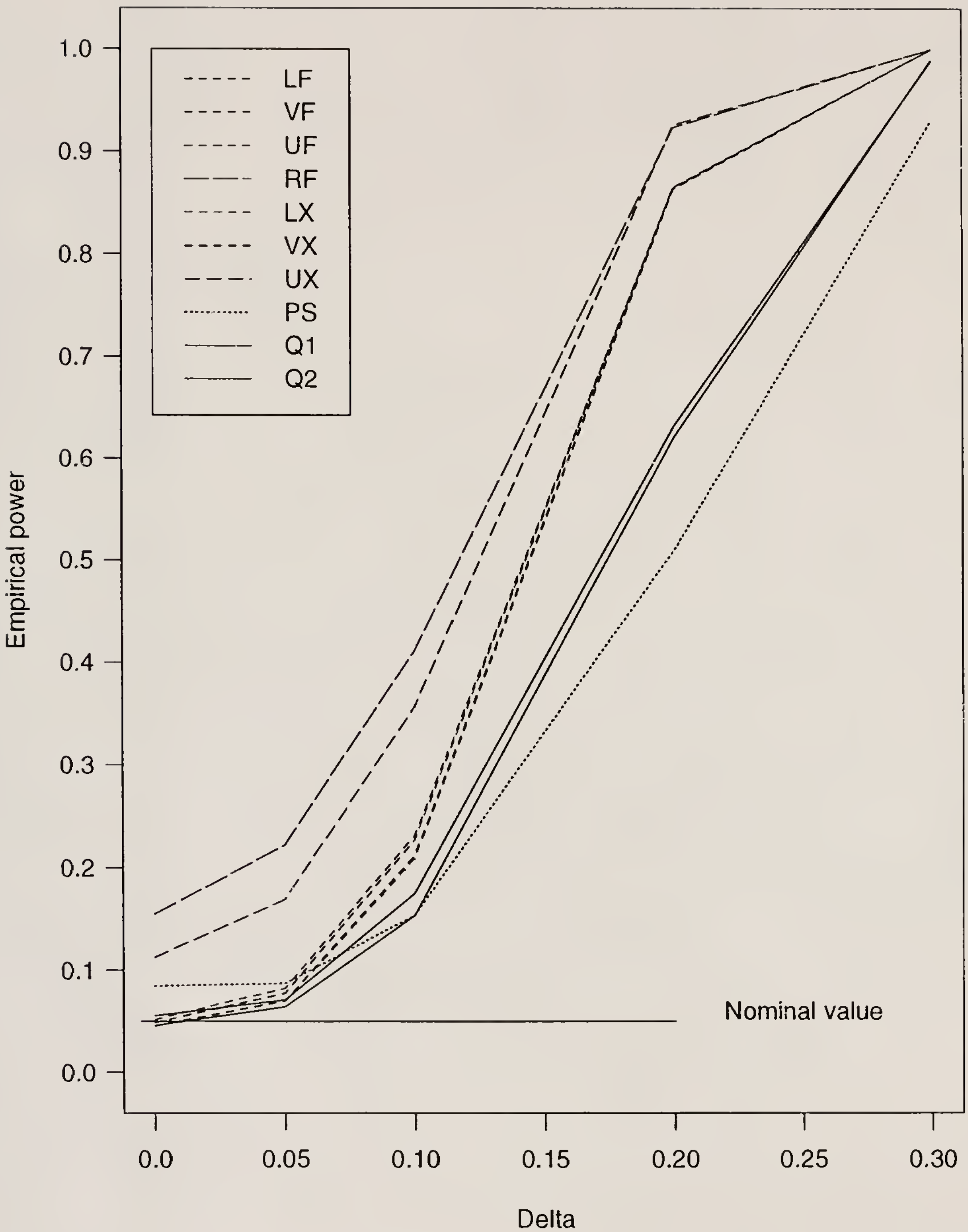

Figure 4.9. $r = 2$, $n = 30$, $\nu = 1$, reps = 2500

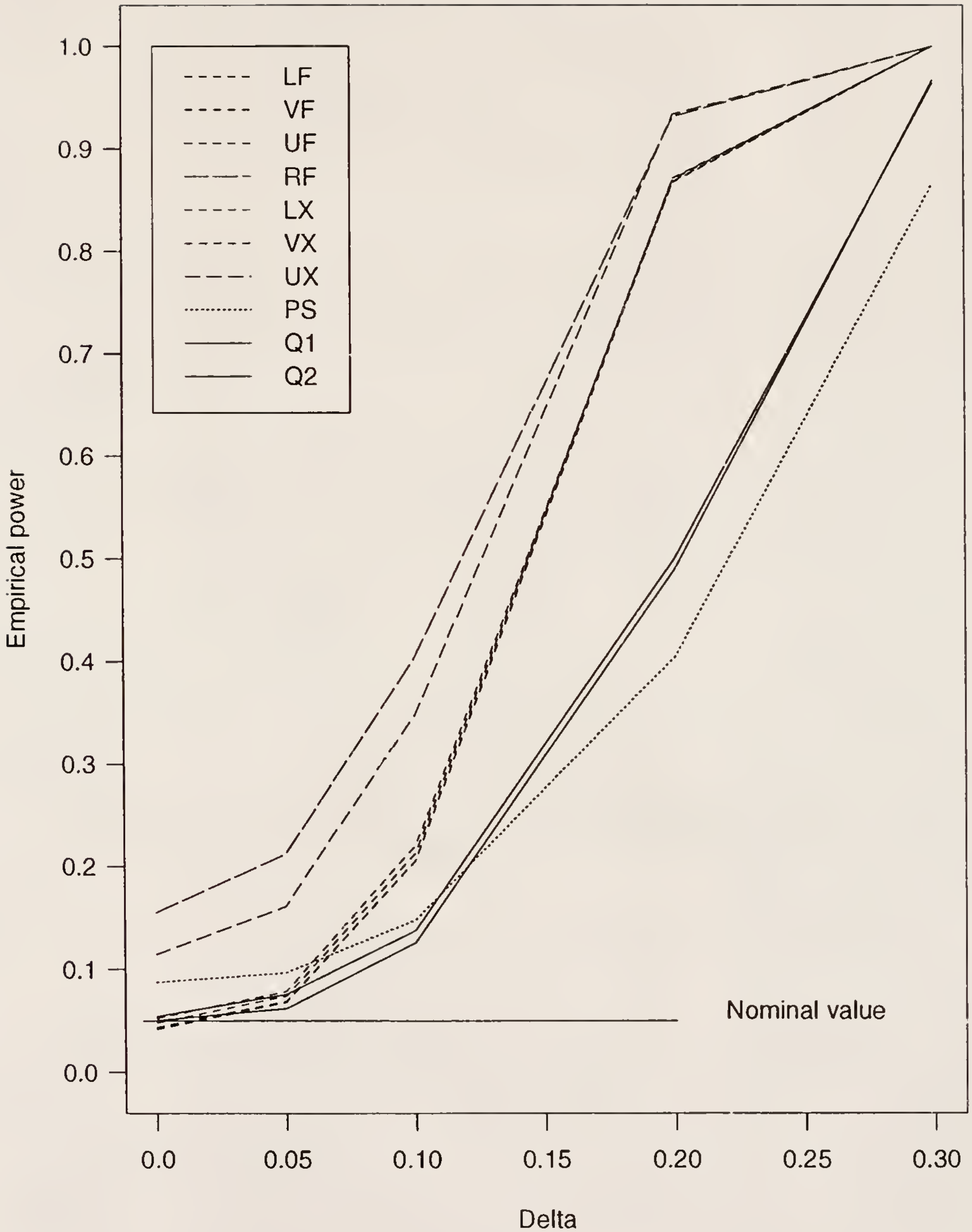

Figure 4.10. $r = 2$, $n = 30$, $\nu = 10$, reps $= 2500$

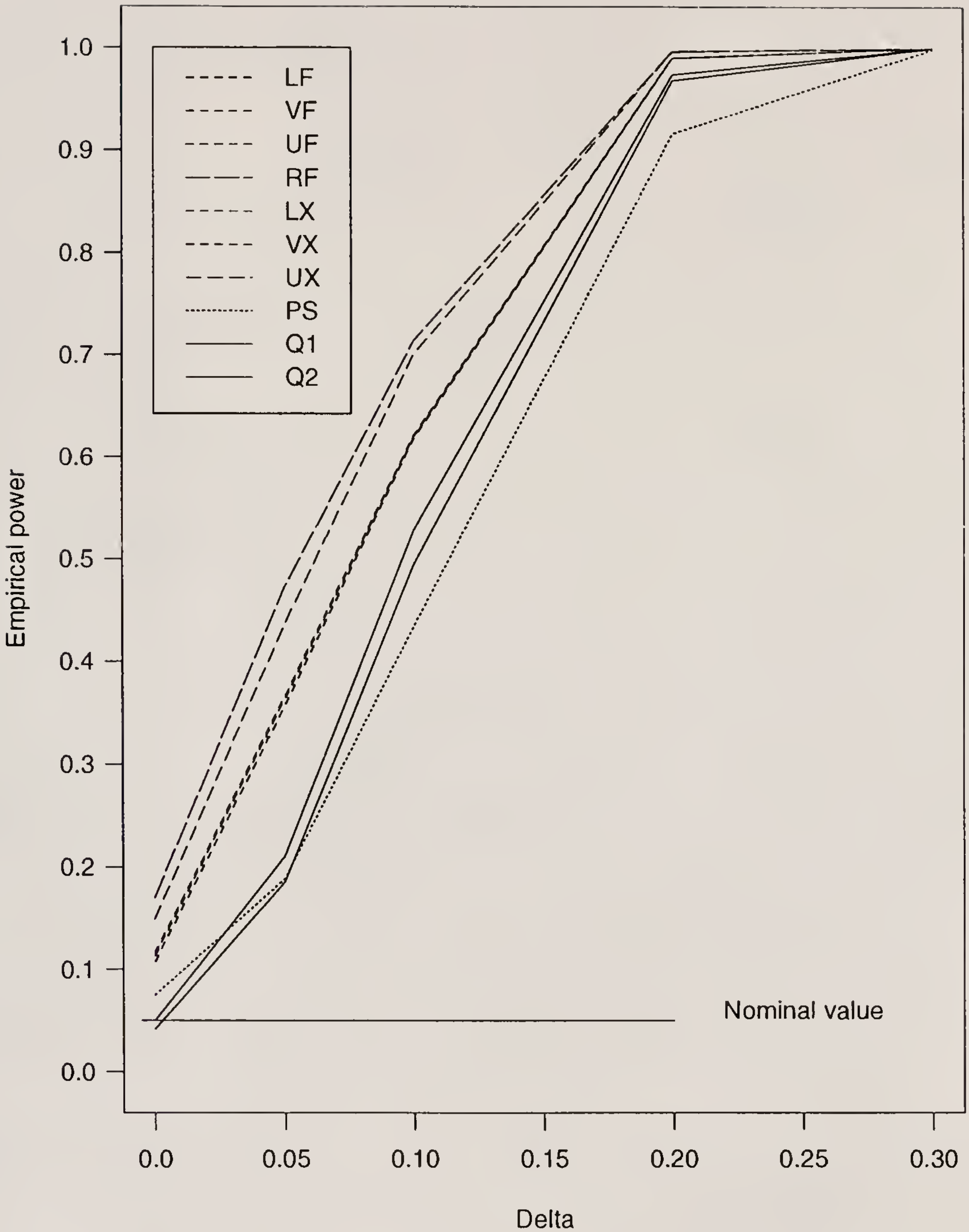

Figure 4.11. $r = 2$, $n = 30$, $df = 1$, reps = 2500

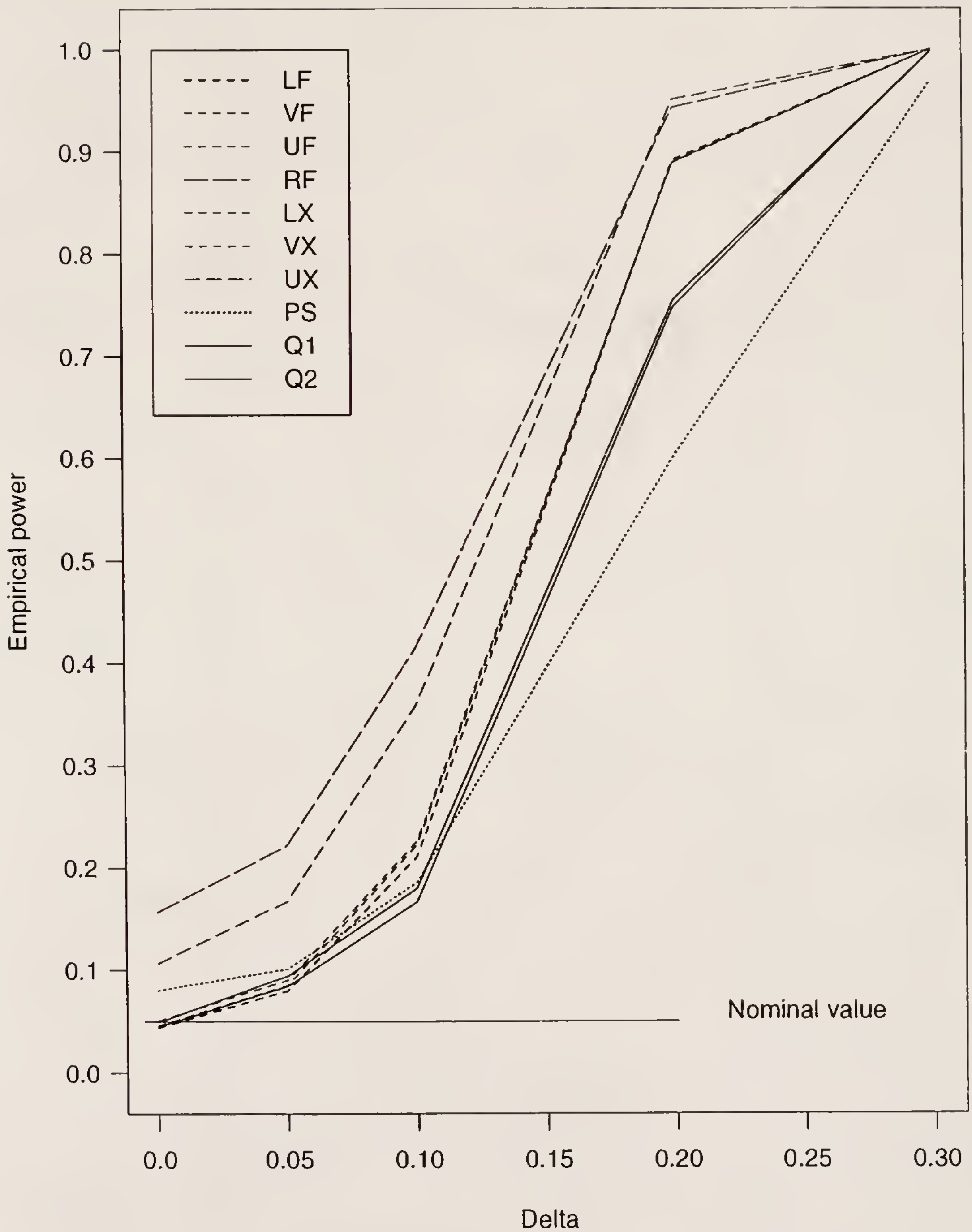

Figure 4.12. $r = 2$, $n = 30$, $df = 5$, reps $= 2500$

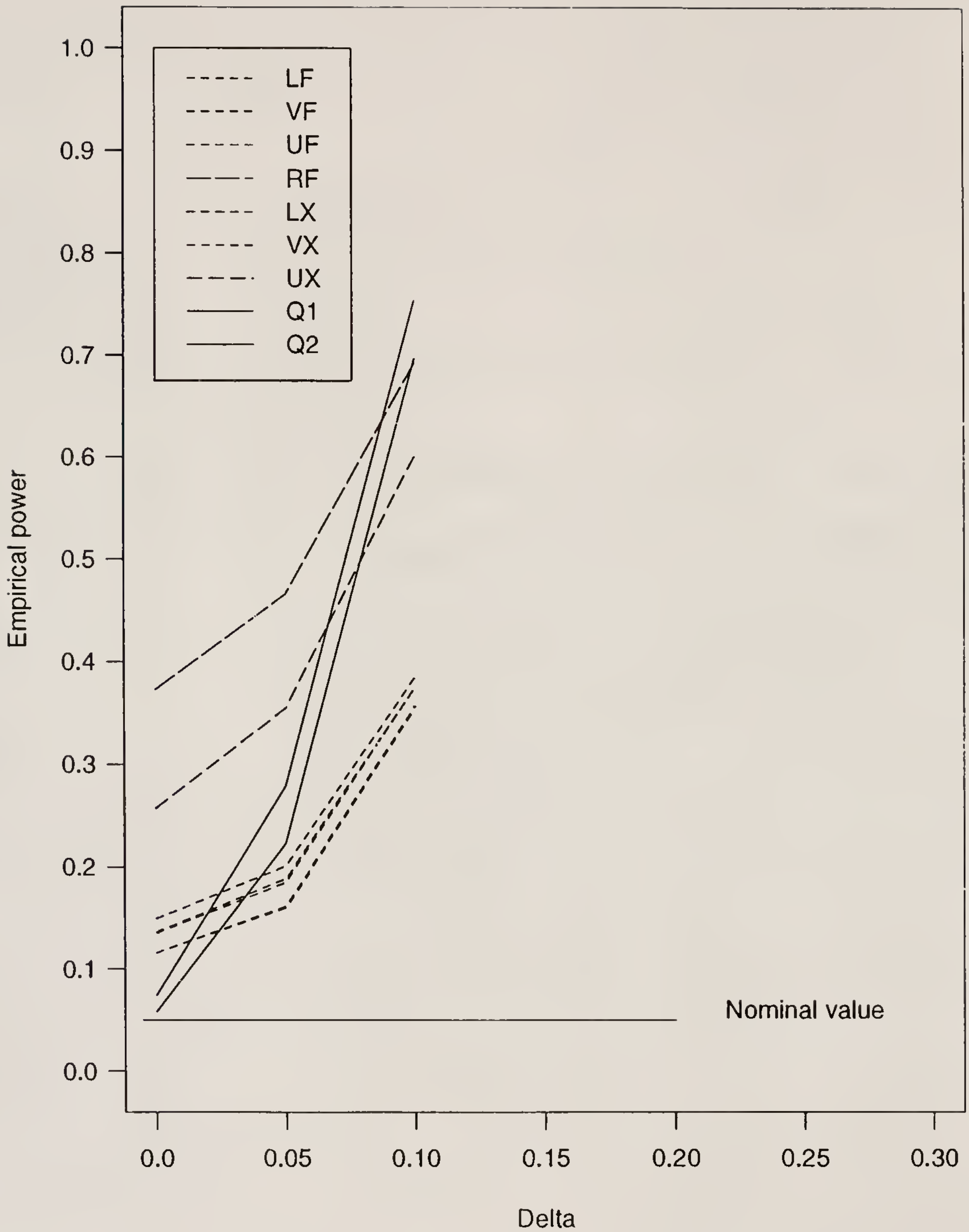

Figure 4.13. $r = 3$, $n = 30$, $\nu = 0.1$, reps = 1000

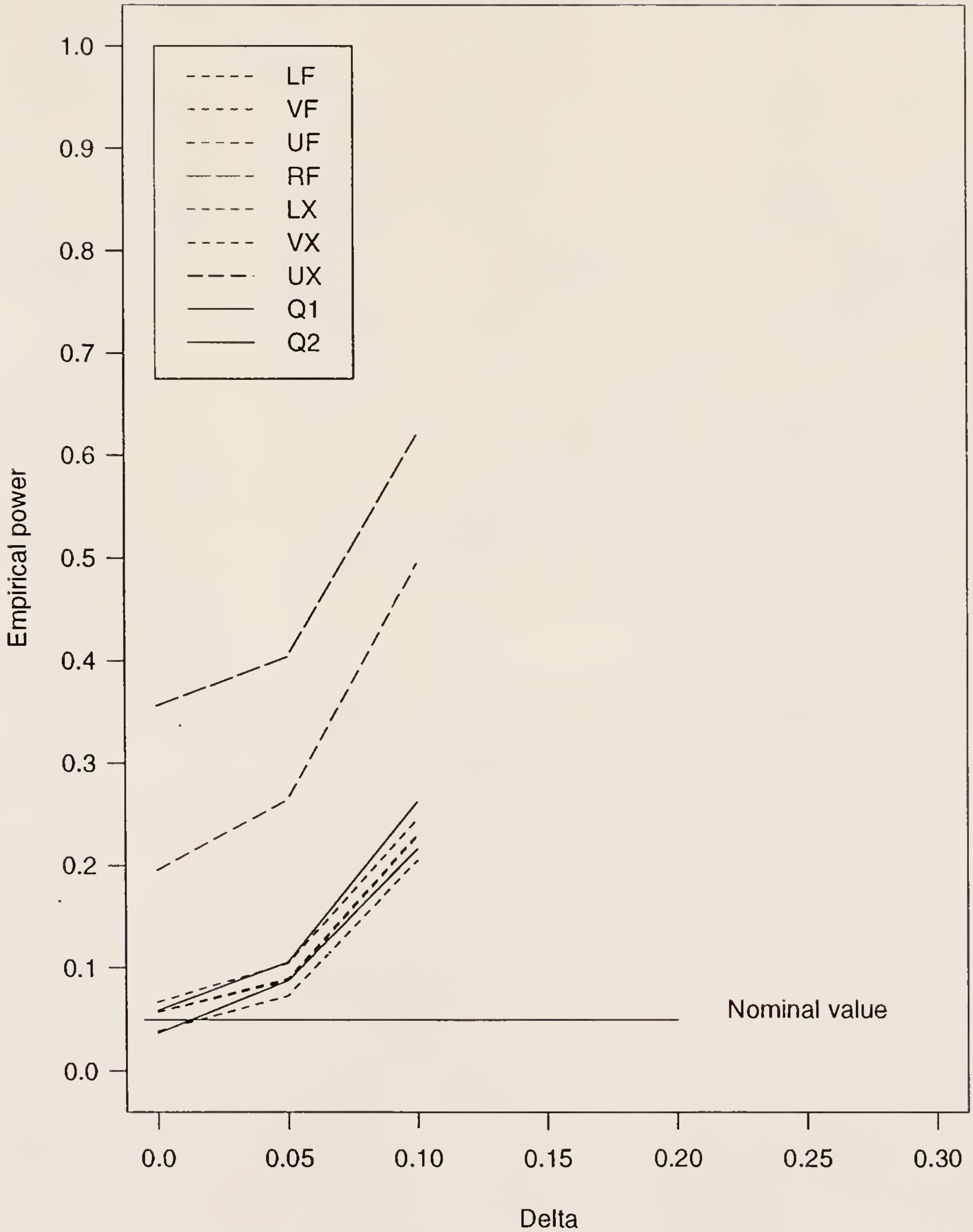

Figure 4.14. $r = 3$, $n = 30$, $\nu = 0.5$, reps = 1000

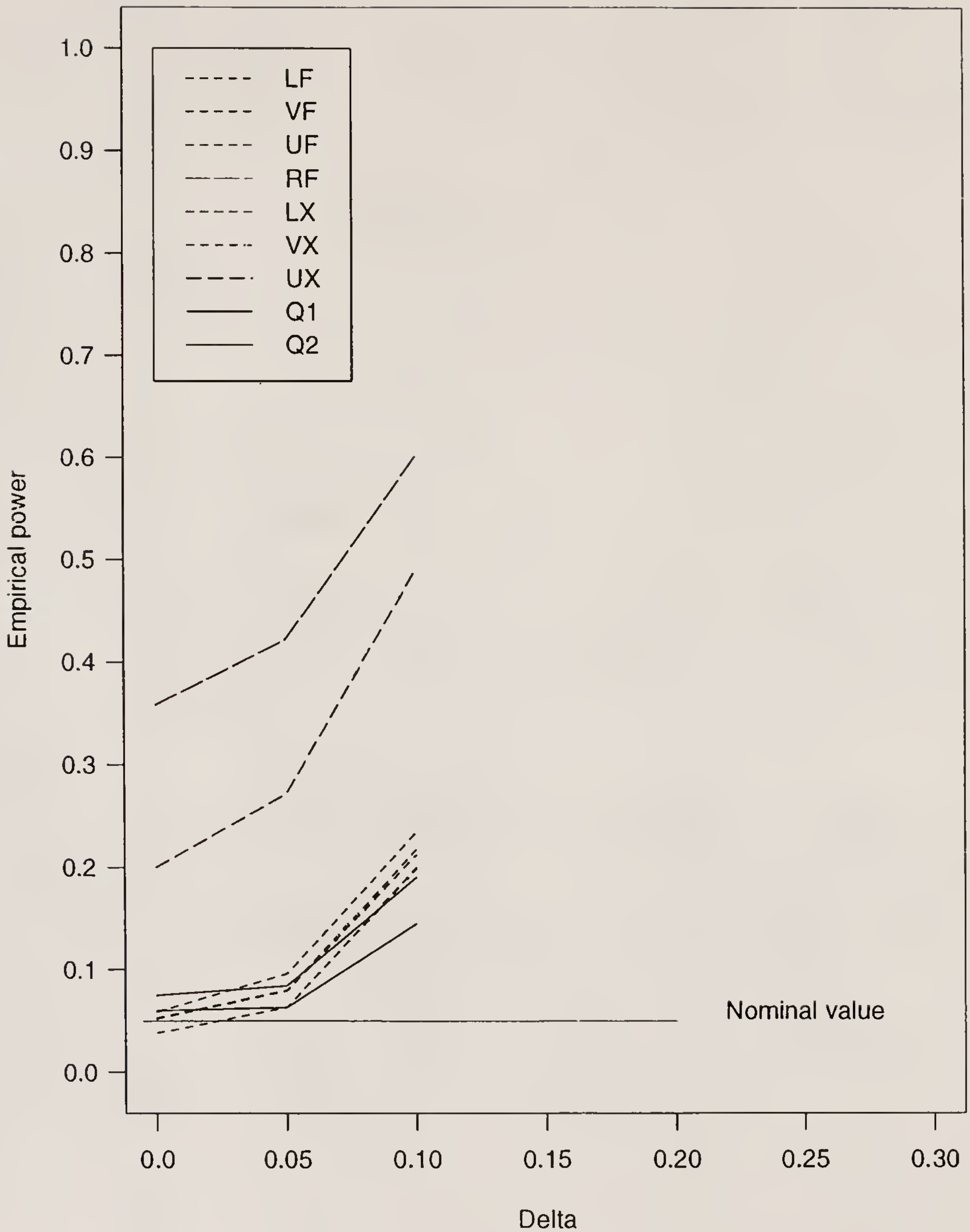

Figure 4.15. $r = 3$, $n = 30$, $\nu = 1$, reps = 1000

CHAPTER 5
APPLICATIONS

5.1 Analysis of Newborn Blood Gas Data

The State of Florida mandates that measurements be taken of newborn blood at birth. This is usually done from the umbilical cord, but it is unclear whether practitioners routinely get umbilical arterial blood or umbilical venous blood. It is further unclear whether these values are related to a simple arterial draw from the abdomen that can be done one hour later in the nursery. Behnke, Eyler, Conlon, Woods, and Thomas (1993) investigate the use of birth weight, gestational age, Apgar scores, cord blood gas values, and first arterial blood gas values as diagnostic criteria for perinatal asphyxia and subsequent low neurodevelopmental outcome in very low birth weight infants. Their data consist of 57 cases for which good blood gas information exists at all three sites—umbilical venous at birth (UV), umbilical arterial at birth (UA), and abdominal arterial one hour after birth (AA). For each site we consider three of the blood gas measurements taken: bicarbonate (HCO_3), partial pressure of carbon dioxide (CO_2), and partial pressure of oxygen (O_2).

We are interested in determining if there is any relationship between the UA and UV blood draws at birth and also if either of these are related to the AA blood draw. Intuitively, since the UA blood originates from the mother and the UV blood from the newborn, we expect that there should not be as much dependence between UV and UA, and UV and AA blood, as there is between UA and AA blood. If these conjectures are supported by the data, there is need for concern with respect to the methods of the practitioners. Practically speaking, the goal of mandatory blood

draws on newborn babies when the diagnostic potential of the measurements may well depend on the technique used seems a flawed idea at best.

A summary of the tests of independence is in Table 5.2. Notice that all the normal theory tests detect dependence in each of the three comparisons, while the nonparametric tests only indicate dependence between the UA and AA measurements. In fact, there is strong agreement among all tests that there is dependence between the UA and AA blood gas measurements. A visual inspection of the blood gas measurements however (see Figures 5.1, 5.2, and 5.3 for plots created by `xgobi`—a program which allows three-dimensional rotation of the data), reveals three observations in the UA determinations which are quite distant from the remainder of the data.

After deleting these observations (28, 52, and 55), the tests were re-run. The results are fairly telling. The normal normal theory tests involving the UV blood gas measurements are now non-significant, with the exception of RF. Of particular interest is how drastically the p-values of these tests changed with the deletion of just three points out of 57. This would seem to be an undesirable property of any test. Although PS managed the same conclusions as Q1 and Q2 (with and without the outliers), its p-value also changed significantly making it suspect also. Only Q1 and Q2 remained largely unaffected by the removal of the outliers.

Although this data is useful in illustrating $\hat{Q}_n$'s resistance to outliers, it is deficient in that there is structure present which is ignored by all these statistics. HCO_3, CO_2, and O_2 are repeated in each set of variates, but because of the invariance of the statistics with respect to the labeling of the variables within a set they fail to take advantage of this fact. Thus, there are probably other techniques for analyzing this data which might be more powerful. The next section illustrates an example where there is no such natural pairing.

Table 5.1: Newborn Blood Data

Obs	UV			UA			AA		
	HCO_3	CO_2	O_2	HCO_3	CO_2	O_2	HCO_3	CO_2	O_2
1	22.8	37	246	22.4	40	34	23.9	40	36
2	19.8	28	80	16.4	39	28	18.4	36	43
3	24.2	45	44	22.0	55	9	27.0	53	15
4	16.9	18	171	23.3	48	21	21.8	39	31
5	13.0	19	83	22.3	46	9	21.2	37	23
6	14.1	26	70	22.4	46	21	21.5	40	30
7	24.2	48	40	19.3	39	29	22.0	42	30
8	17.6	39	81	21.8	34	26	21.1	29	41
9	21.8	68	34	15.6	51	33	19.3	44	32
10	18.4	99	67	22.1	43	16	20.3	28	36
11	16.4	23	110	23.5	50	19	22.8	40	25
12	19.8	21	63	22.1	47	18	23.2	38	27
13	22.8	42	73	23.2	48	39	23.0	44	52
14	17.2	16	365	21.4	42	31	22.8	41	37
15	13.2	21	280	16.9	43	17	19.0	35	27
16	23.9	94	66	19.7	42	13	19.0	34	26
17	20.6	34	169	25.7	54	12	22.3	43	23
18	22.5	29	207	13.5	27	30	22.5	42	39
19	15.1	25	66	17.0	62	23	20.6	39	64
20	16.8	21	434	25.4	47	20	22.9	39	51
21	24.8	63	45	26.0	58	14	25.0	49	23
22	23.0	26	49	21.8	40	21	22.9	38	29
23	19.5	30	74	24.8	50	21	24.4	45	25
24	20.9	39	39	22.2	42	14	20.0	35	20
25	19.6	38	48	24.9	45	17	23.3	38	29
26	22.1	53	48	24.5	42	20	23.2	36	31
27	17.0	28	85	27.7	70	7	24.8	44	30
28	16.8	29	47	21.0	111	3	20.3	103	6
29	19.6	34	194	24.5	57	11	26.4	50	27
30	17.0	26	283	19.5	44	20	18.6	33	28
31	18.2	18	64	16.6	56	21	21.2	47	31
32	22.4	20	271	20.5	38	22	21.3	35	32
33	27.0	49	36	26.6	50	24	25.5	45	34
34	21.5	50	56	22.8	51	10	20.6	42	24
35	19.7	58	35	23.0	46	20	22.4	40	38
36	23.1	41	225	24.7	45	22	22.2	37	27
37	22.2	47	60	22.3	58	19	20.7	45	27
38	18.8	21	173	21.6	51	14	25.3	49	25
39	11.4	9	87	21.2	46	9	26.2	54	14
40	21.1	36	131	19.5	37	30	19.3	35	28
41	23.6	47	198	21.9	45	10	21.8	42	15

Table 5.1: —continued

Obs	UV			UA			AA		
	HCO_3	CO_2	O_2	HCO_3	CO_2	O_2	HCO_3	CO_2	O_2
42	18.8	44	46	20.3	57	13	21.0	53	19
43	12.3	12	300	18.7	45	15	20.2	41	15
44	13.9	12	270	18.7	45	8	22.8	50	22
45	21.9	42	194	19.7	58	23	19.6	53	26
46	16.6	27	27	20.4	42	30	19.6	38	70
47	19.4	27	70	22.7	42	16	22.4	34	23
48	21.3	28	38	23.9	49	15	23.9	46	21
49	20.8	44	299	21.9	43	24	18.7	32	40
50	21.4	42	163	24.3	51	16	23.3	44	25
51	19.6	38	199	21.6	41	29	21.3	33	40
52	26.2	37	135	25.6	23	51	25.6	44	56
53	16.1	26	40	21.1	35	25	20.2	30	35
54	19.5	37	194	22.2	56	14	19.4	42	26
55	11.1	27	42	16.9	115	11	17.2	96	18
56	15.2	25	94	19.6	46	19	18.6	32	28
57	22.2	50	76	21.3	38	25	22.0	38	28

5.2 Analysis of Fitness Club Data

The section of SAS/STAT User's Guide, Volume 1, which describes the SAS procedure PROC CANCORR has an example that uses data provided by Dr. A. C. Linnerud, North Carolina State University, in which three physiological variables and three exercise variables were measured on twenty middle-aged men in a fitness club. It demonstrates how PROC CANCORR can be used to determine if the physiological variables are related in any way to the exercise variables.

Here again we have potential outliers, which we can see from examining Figures 5.4 and 5.5. A summary of the statistical analysis of this data (with and without observations 10 and 14) is in Table 5.4. Interestingly, the p-values for the normal tests all decrease while the p-values for the nonparametric tests all increase with the removal of the outliers.

Table 5.2. Statistical Analysis of Newborn Blood Data

	UV vs. UA		UV vs. AA		UA vs. AA	
Statistic	Value	P-value	Value	P-value	Value	P-value
LF	2.5649	0.0097	2.1290	0.0317	35.6598	0.0000
VF	2.3596	0.0158	2.0662	0.0356	22.0281	0.0000
UF	2.7189	0.0058	2.1519	0.0285	45.9539	0.0000
RF	7.8735	0.0002	5.2465	0.0030	122.2455	0.0000
PS	12.8302	0.1704	7.0634	0.6305	70.6444	0.0000
Q1	4.9852	0.8356	9.6629	0.3785	82.9273	0.0000
Q2	5.2102	0.8156	8.1707	0.5170	82.5464	0.0000
after the outliers are removed						
LF	1.1296	0.3476	1.0216	0.4271	14.7553	0.0000
VF	1.1137	0.3564	0.9938	0.4477	12.8989	0.0000
UF	1.1386	0.3397	1.0446	0.4080	14.5733	0.0000
RF	3.2089	0.0308	3.2942	0.0279	29.3290	0.0000
PS	6.4742	0.6917	4.9757	0.8364	59.2939	0.0000
Q1	3.2630	0.9530	9.4586	0.3961	75.7127	0.0000
Q2	3.8196	0.9229	7.4735	0.5880	75.4526	0.0000

5.3 Analysis of Cotton Dust Data

Merchant et al. (1975) studied the effects of cotton dust exposure on human beings by measuring several respiratory variables and several blood-related variables on 12 subjects exposed for six hours. The data consist of changes in these variables from baseline. It may be of medical interest to determine if these two sets of variables are independent or not. Included among the respiratory variables are closing capacity (CC), vital capacity (VC), and total lung capacity (TLC). Two blood-related variables are oxygen (O_2) and white blood count (WBC).

Table 5.3. Fitness Club Data

Obs.	Weight	Waist	Pulse	Chinups	Situps	Jumps
1	191	36	50	5	162	60
2	189	37	52	2	110	60
3	193	38	58	12	101	101
4	162	35	62	12	105	37
5	189	35	46	13	155	58
6	182	36	56	4	101	42
7	211	38	56	8	101	38
8	167	34	60	6	125	40
9	176	31	74	15	200	40
10	154	33	56	17	251	250
11	169	34	50	17	120	38
12	166	33	52	13	210	115
13	154	34	64	14	215	105
14	247	46	50	1	50	50
15	193	36	46	6	70	31
16	202	37	62	12	210	120
17	176	37	54	4	60	25
18	157	32	52	11	230	80
19	156	33	54	15	225	73
20	138	33	68	2	110	43

Table 5.4. Statistical Analysis of Fitness Club Data

	Outliers In		Outliers Out	
Statistic	Value	P-value	Value	P-value
LF	2.0482	0.0638	2.4847	0.0305
VF	1.5587	0.1551	1.5877	0.1504
UF	2.4938	0.0238	3.4540	0.0045
RF	9.1986	0.0009	13.5137	0.0002
PS	NA	NA	15.4118	0.0802
Q1	12.9838	0.1633	11.8387	0.2226
Q2	13.4917	0.1416	12.5776	0.1827

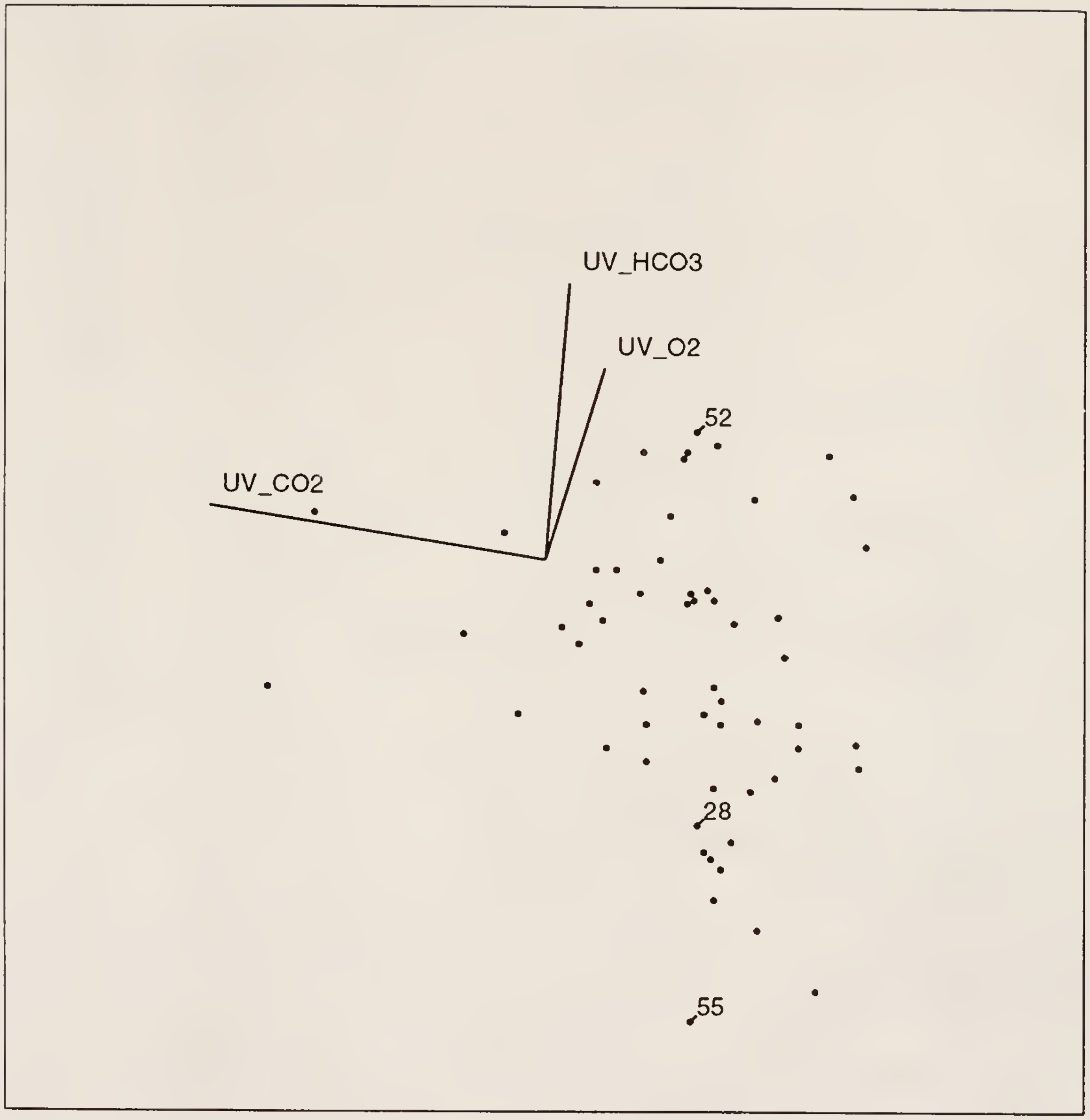

Figure 5.1. Umbilical Venous Blood Gas Measurements

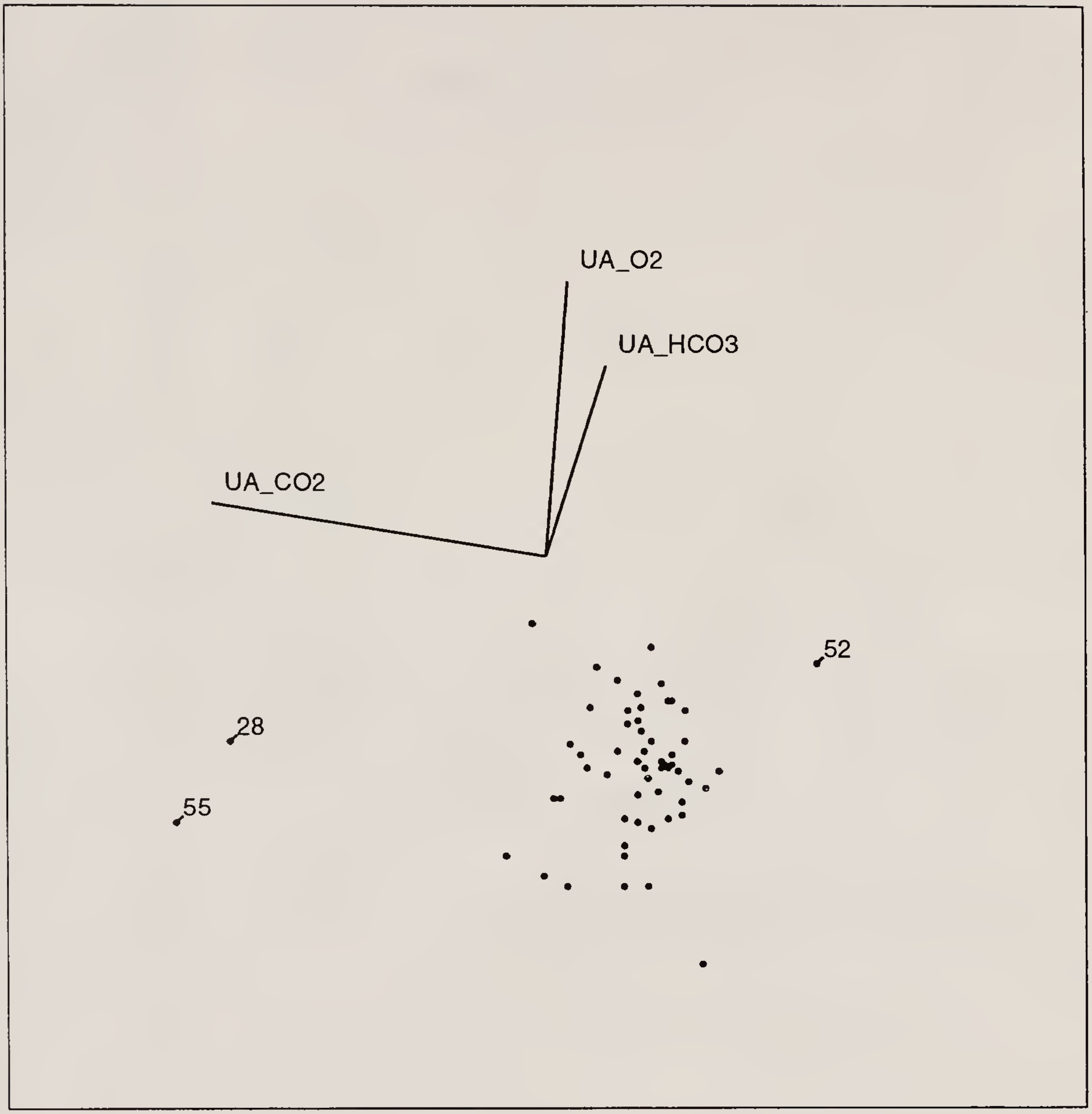

Figure 5.2. Umbilical Arterial Blood Gas Measurements

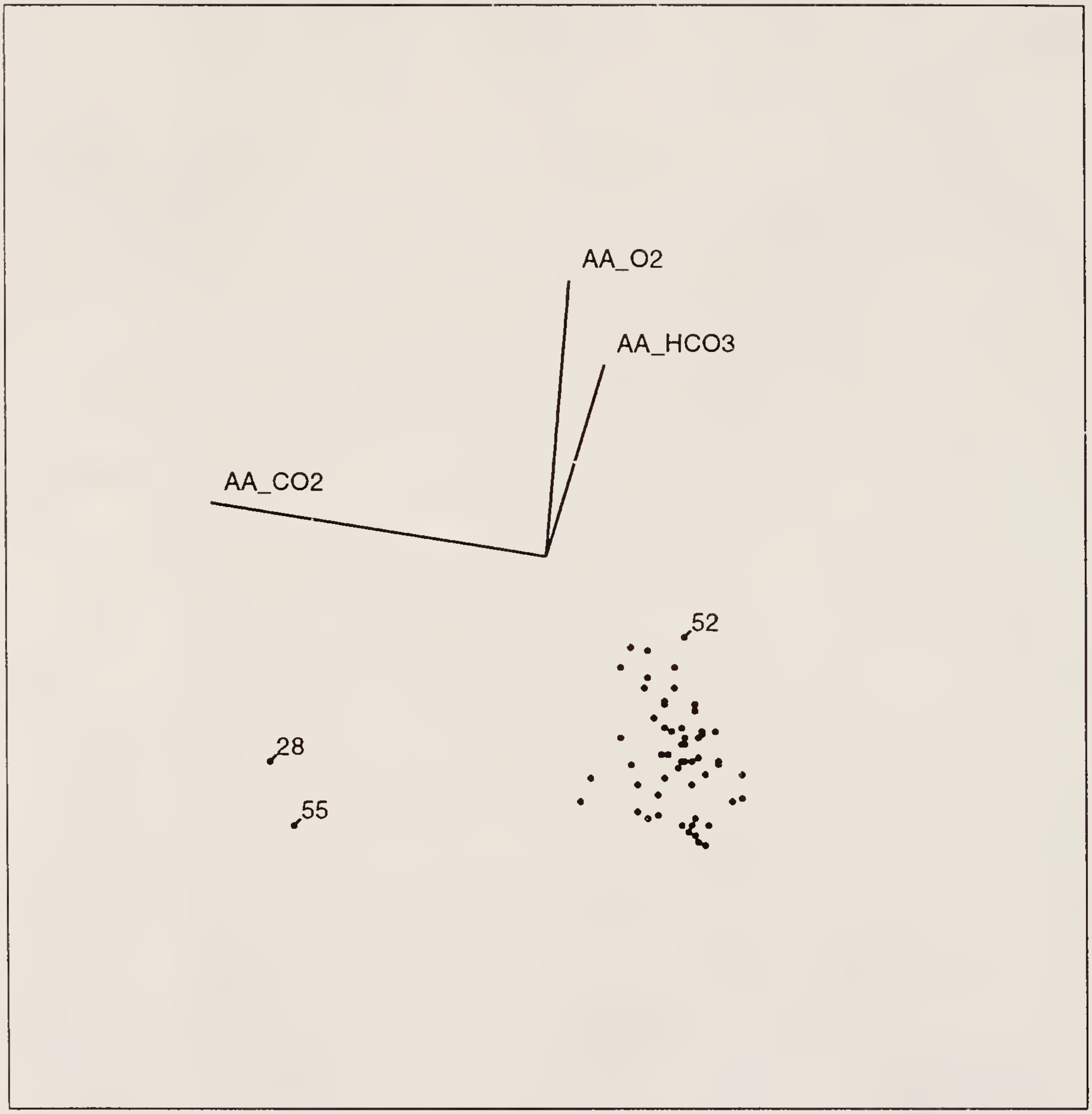

Figure 5.3. Abdominal Arterial Blood Gas Measurements

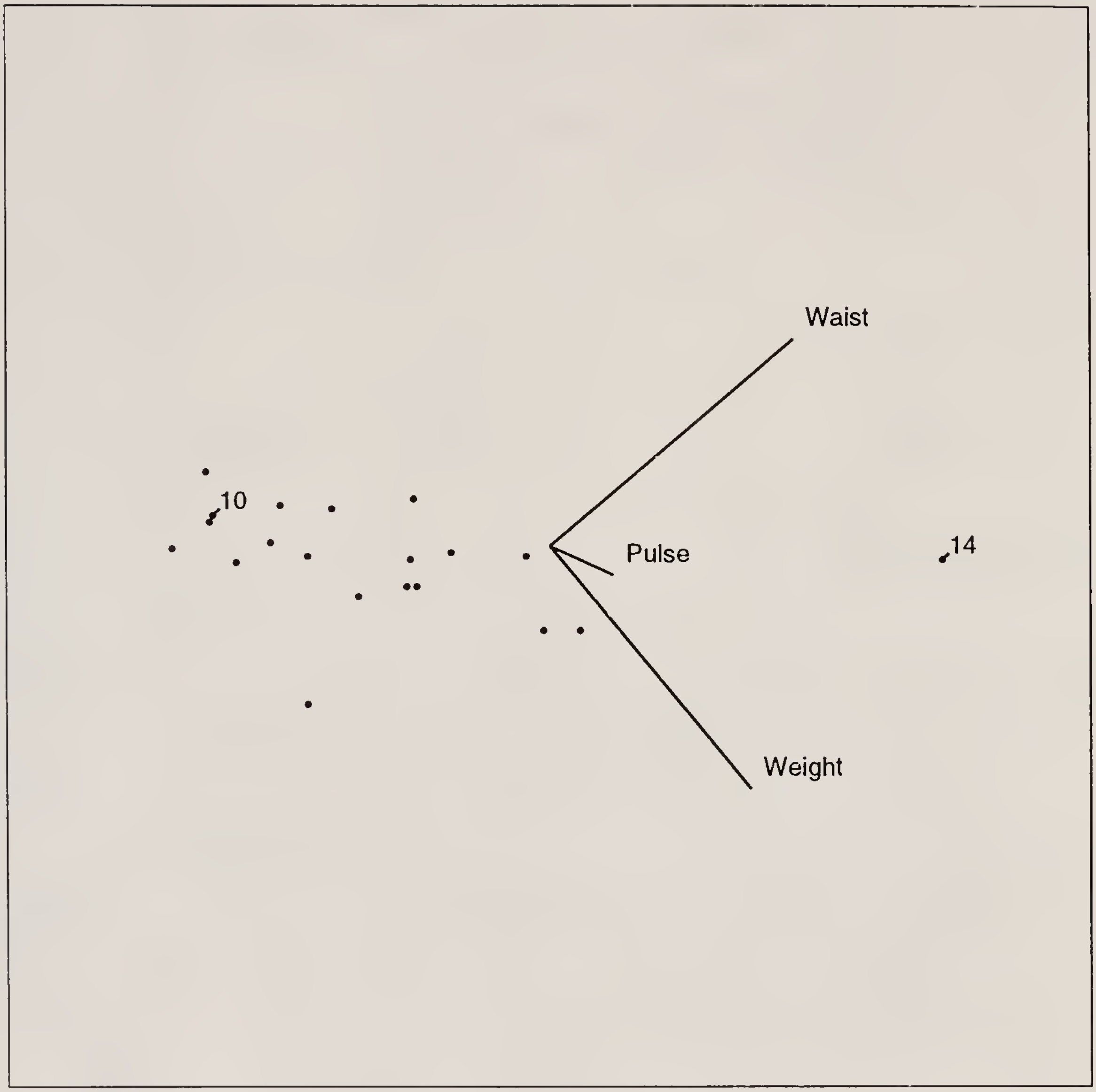

Figure 5.4. Physiological Measurements

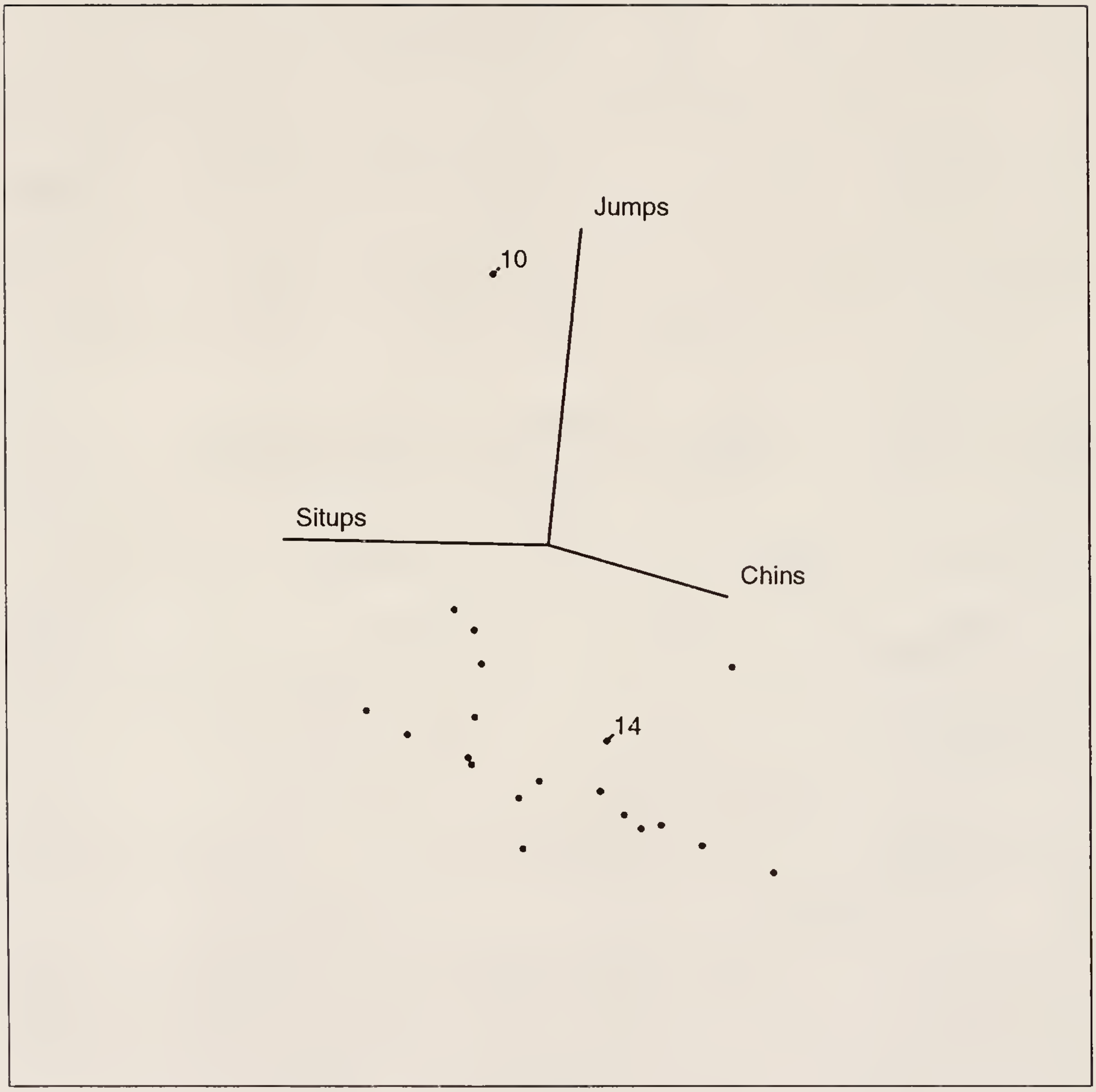

Figure 5.5. Exercise Measurements

Table 5.5. Cotton Dust Data

Obs.	CC	VC	TLC	O_2	WBC
1	-4.3	0.24	-0.11	-1.0	6000
2	4.4	-0.29	-0.01	-1.5	-350
3	7.5	0.10	0.67	1.0	-250
4	-0.3	0.13	0.31	6.5	1675
5	-5.8	0.02	-0.75	-3.0	875
6	14.5	0.48	1.14	-3.0	-100
7	-1.9	-0.05	-0.22	-15.5	1075
8	17.3	-0.62	0.62	-13.5	1675
9	2.5	-0.16	0.12	0.0	1500
10	-5.6	0.15	-0.14	-4.0	2200
11	2.2	0.25	0.40	-2.5	650
12	5.5	-0.42	0.22	-1.0	3025

Table 5.6. Statistical Analysis of Cotton Dust Data

Statistic	Value	P-value
LF	0.8623	0.5452
VF	0.9339	0.4976
UF	0.7783	0.6024
RF	1.6581	0.2521
PS	2.4917	0.8694
Q1	14.7390	0.0224
Q2	13.4729	0.0361

CHAPTER 6
CONCLUSION

6.1 Discussion

With the increasing availability of very fast computers, multivariate and nonparametric procedures which would have been impossible for the average statistician to implement several years ago can easily be utilized today. This fact is responsible for the explosion of such multivariate and nonparametric methodologies seen currently. The goal of this dissertation has been to research one such methodology, the interdirection quadrant statistic ($\hat{Q}_n$), in testing for independence between two sets of variates. The basic competitors considered were Wilks' likelihood ratio criterion ($-n \log V$) and a specific member of a class of statistics invented by Puri and Sen ($-n \log S^{J_0}$). As demonstrated in Chapter 3, the Pitman ARE's indicate that $\hat{Q}_n$ does quite well relative to $-n \log V$ for heavy-tailed distributions and is competitive for moderate-tailed distributions. The statistic $\hat{Q}_n$, under spherical alternatives, appears to be uniformly better than its natural nonparametric competitor $-n \log S^{J_0}$. Simulation results concur with theoretical findings in the sense that the empirical powers of the competitors are ordered in the same way as the Pitman ARE's indicate they should be.

6.2 Further Research

Other avenues of research might be to investigate the potential of the interdirection quadrant statistic (i) for describing the nature of the association between two sets

of variates instead of merely using it as a test of independence, or (ii) for determining which variables in a set may or may not be contributing to an association between the two sets. It might also be desirable to make comparisons with other members of Puri and Sen's class of statistics, like a multivariate analog of Spearman's rho. Computing issues that arose during the simulation study also lead to possible research areas. Since the time to compute the interdirections (based on simple looping algorithms) is on the order of n^k, where k is the dimension, for even moderate sample sizes, use of the interdirections becomes impractical. Possible work-arounds to this limitation might be (i) to find a suitable approximating statistic (which might entail more simulation work), (ii) to derive a faster algorithm for computing the interdirections, or possibly (iii) to estimate, by using some sampling method for instance, the interdirections. Of course, since the interdirections have wider applicability than the present independence testing situation, these results would naturally have broader appeal than this context. The statistic Q2 defined in Chapter 4 and included in the simulations there, is an example of a preliminary step in examining the feasibility of (i). It appears to do quite well, so that this seems to be a promising idea.

APPENDIX A
CONVERGENCE RESULTS

A recurring theme in the proofs of the lemmas to follow is the re-expression of certain terms arising from Taylor series expansions as U-statistics. This is important in order to apply existing U-statistic theory to these quantities. Specifically, we will have occasion to use Theorem 2.8 in Randles (1982), which we restate here. Randles considers random variables which would be U-statistics were it not for the fact that they contain an estimator. Let $X_1, \ldots, X_n$ denote a random sample from some population. Let $h(x_1, \ldots, x_r; \gamma)$ denote a symmetric kernel of degree r with expected value $\mu(\gamma) = E_\lambda[h(X_1, \ldots, X_r; \gamma)]$, where λ denotes the true value of an unknown parameter γ. Let $U_n(\gamma) = \binom{n}{r}^{-1} \sum_{\alpha \in A^*} h(X_{\alpha_1}, \ldots, X_{\alpha_r}; \gamma)$, where A^* denotes the collection of all subsets of size r from the integers $\{1, \ldots, n\}$. Consider the following conditions:

Condition 1. Suppose

$$(\hat{\lambda} - \lambda) = O_p(n^{-1/2}).$$

Condition 2. Suppose there is a neighborhood of λ, call it $K(\lambda)$, and a constant $K_1 > 0$, such that if $\gamma \in K(\lambda)$ and $D(\gamma, d)$ is a sphere centered at γ with radius d satisfying $D(\gamma, d) \subset K(\lambda)$, then

$$E\left[\sup_{\gamma' \in D(\gamma,d)} \left|h(X_1, \ldots, X_r; \gamma') - h(X_1, \ldots, X_r; \gamma)\right|\right] \leq K_1 d.$$

Condition 3. Suppose there exists an $M_1 > 0$ such that

$$|h(x_1, \ldots, x_r; \gamma) - h(x_1, \ldots, x_r; \lambda)| \leq M_1$$

for every $(x_1, \ldots, x_r)$ and all γ in some neighborhood of λ.

Theorem A.0.1 If Conditions 1, 2, and 3 are satisfied, then

$$U_n(\hat{\lambda}) - \mu(\hat{\lambda}) - U_n(\lambda) + \mu(\lambda) = o_p(n^{-1/2}).$$

Theorem A.0.1 also holds in the multivariate setting, so we can readily apply it to the present situation. Also, since all the kernels we will work with are bounded, and $(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1) = O_p(n^{-1/2})$, we will only be concerned with satisfying Condition 2 in Theorem A.0.1. In the lemmas that follow, the sufficient conditions will not be given in the statement of the lemma since they are quite technical and the lemmas are not of importance in and of themselves. Rather, the relevant conditions will be noted as they are needed. We now prove the first of the necessary limiting results.

Lemma A.0.1 $\bar{c} = o_p(n^{-1/2})$.

Proof of Lemma A.0.1 Define

$$\hat{V}_n(\gamma_1) = \binom{n}{2}^{-1} \sum_{i<j}^{n} \cos(\pi \hat{p}_1(i,j;\gamma_1)),$$

where

$$\hat{p}_1(i,j;\gamma_1) = \binom{n-2}{r_1-1}^{-1} \sum_{\substack{\boldsymbol{\alpha} \in C_n^{r_1-1} \\ (i,j) \notin \boldsymbol{\alpha}}} h_1(\boldsymbol{X}_i^{(1)}, \boldsymbol{X}_j^{(1)}; \boldsymbol{X}_{\alpha_1}^{(1)}, \ldots, \boldsymbol{X}_{\alpha_{r_1-1}}^{(1)}; \gamma_1),$$

and $h_1(\boldsymbol{X}_i^{(1)}, \boldsymbol{X}_j^{(1)}; \boldsymbol{X}_{\alpha_1}^{(1)}, \ldots, \boldsymbol{X}_{\alpha_{r_1-1}}^{(1)}; \gamma_1)$ indicates whether (1) or not (0) $\boldsymbol{X}_i^{(1)}$ and $\boldsymbol{X}_j^{(1)}$ are on opposite sides of the hyperplane formed by $\boldsymbol{X}_{\alpha_1}^{(1)}, \ldots, \boldsymbol{X}_{\alpha_{r_1-1}}^{(1)}$ and γ_1. Since $\bar{c} = \hat{V}(\hat{\boldsymbol{\theta}}_1)$, our immediate goal is to express bounds on $\hat{V}(\boldsymbol{\gamma}_1)$ in terms of U-statistics so that we can apply Theorem A.0.1. Using a Taylor series expansion of $\cos x$ around the point y, ($\cos x = \cos y - (x-y)\sin y - (1/2)(x-y)^2 \cos z$ for some

z between x and y), we have the following representation:

$$\begin{aligned}\hat{V}_n(\gamma_1) &= \binom{n}{2}^{-1} \sum_{i<j}^{n} \cos(\pi p_1(i,j;\gamma_1)) \\ &- \pi \binom{n}{2}^{-1} \sum_{i<j}^{n} \left\{\hat{p}_1(i,j;\gamma_1) - p_1(i,j;\gamma_1)\right\} \sin(\pi p_1(i,j;\gamma_1)) \\ &- \frac{\pi^2}{2}\binom{n}{2}^{-1} \sum_{i<j}^{n} \left\{\hat{p}_1(i,j;\gamma_1) - p_1(i,j;\gamma_1)\right\}^2 \cos(\pi p^*_{ij}),\end{aligned} \tag{A.1}$$

where $p_1(i,j;\gamma_1) = \mathrm{E}_{H_0}\left[h_1(\boldsymbol{X}_i^{(1)}, \boldsymbol{X}_j^{(1)}; \boldsymbol{X}_{\alpha_1}^{(1)}, \ldots, \boldsymbol{X}_{\alpha_{r_1-1}}^{(1)}; \gamma_1) \mid \boldsymbol{X}_i^{(1)}, \boldsymbol{X}_j^{(1)}\right]$ and p^*_{ij} is a value between $\hat{p}_1(i,j;\gamma_1)$ and $p_1(i,j;\gamma_1)$. If we define

$$V_{1n}(\gamma_1) = \binom{n}{2}^{-1} \sum_{i<j}^{n} \cos(\pi p_1(i,j;\gamma_1)),$$

$$V_{2n}(\gamma_1) = \binom{n}{2}^{-1} \sum_{i<j}^{n} \left\{\hat{p}_1(i,j;\gamma_1) - p_1(i,j;\gamma_1)\right\} \sin(\pi p_1(i,j;\gamma_1)),$$

and

$$V_{3n}(\gamma_1) = \binom{n}{2}^{-1} \sum_{i<j}^{n} \left\{\hat{p}_1(i,j;\gamma_1) - p_1(i,j;\gamma_1)\right\}^2,$$

then it follows from (A.1) that

$$|\hat{V}_n(\gamma_1)| \leq |V_{1n}(\gamma_1)| + \pi|V_{2n}(\gamma_1)| + (\pi^2/2)|V_{3n}(\gamma_1)|. \tag{A.2}$$

Since Randles (1989) has shown that his interdirection sign statistic has a limiting chi-square distribution under the null hypothesis of a known location parameter, in our notation this result is equivalent to $r_1((n-1)V_{1n}(\boldsymbol{\theta}_1) + 1) \xrightarrow{d} \chi^2_{r_1}$. Thus, $V_{1n}(\boldsymbol{\theta}_1)$ is $O_p(n^{-1})$. Applying Theorem A.0.1 and a Taylor series expansion of $\mu_{V_{1n}}(\hat{\boldsymbol{\theta}}_1)$

around $\boldsymbol{\theta}_1$, we get

$$\begin{aligned}
V_{1n}(\hat{\boldsymbol{\theta}}_1) &= V_{1n}(\boldsymbol{\theta}_1) + (\mu_{V_{1n}}(\hat{\boldsymbol{\theta}}_1) - \mu_{V_{1n}}(\boldsymbol{\theta}_1)) + o_p(n^{-1/2}) \\
&= V_{1n}(\boldsymbol{\theta}_1) + (\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1)'\nabla\mu_{V_{1n}}(\boldsymbol{\theta}_1^*) + o_p(n^{-1/2}) \\
&= O_p(n^{-1}) + \boldsymbol{O}_p'(n^{-1/2})\boldsymbol{o}_p(1) + o_p(n^{-1/2}) \\
&= o_p(n^{-1/2}),
\end{aligned}$$

if $\nabla\mu_{V_{1n}}(\boldsymbol{\gamma}_1)$ exists and is continuous at θ_1, since then $\nabla\mu_{V_{1n}}(\boldsymbol{\gamma}_1)|_{\gamma_1=\boldsymbol{\theta}_1} = \mathbf{0}$ (the directional derivative at θ_1 is constant) and $\boldsymbol{\theta}_1^* = \boldsymbol{\theta}_1 + \boldsymbol{o}_p(1)$. We label the existence and continuity condition as (C1). Condition 2 is satisfied for the kernel $\cos(\pi p_1(\boldsymbol{x}_1^{(1)}, \boldsymbol{x}_2^{(1)}; \gamma_1))$ if it is satisfied for the kernel $p_1(\boldsymbol{x}_1^{(1)}, \boldsymbol{x}_2^{(1)}; \gamma_1))$ and we label this condition as (C2).

With a little work, $V_{2n}(\boldsymbol{\gamma}_1)$ is seen to be a multiple of a U-statistic.

$$\begin{aligned}
V_{2n}(\gamma_1) &= \binom{n}{2}^{-1}\binom{n-2}{r_1-1}^{-1}\sum_{i<j}^{n}\sum_{\substack{\boldsymbol{\alpha}\in C_n^{r_1-1} \\ (i,j)\notin\boldsymbol{\alpha}}} g_2(\boldsymbol{X}_i^{(1)}, \boldsymbol{X}_j^{(1)}; \boldsymbol{X}_{\alpha_1}^{(1)}, \ldots, \boldsymbol{X}_{\alpha_{r_1-1}}^{(1)}; \gamma_1) \\
&= \binom{r_1+1}{2}^{-1}\binom{n}{r_1+1}^{-1}\sum_{\boldsymbol{\beta}\in C_n^{r_1+1}} g_2^*(\boldsymbol{X}_{\beta_1}^{(1)}, \ldots, \boldsymbol{X}_{\beta_{r_1+1}}^{(1)}; \gamma_1) \\
&= \binom{r_1+1}{2}^{-1} U_{2n}(\gamma_1),
\end{aligned}$$

where

$$\begin{aligned}
&g_2(\boldsymbol{X}_1^{(1)}, \boldsymbol{X}_2^{(1)}; \boldsymbol{X}_3^{(1)}, \ldots, \boldsymbol{X}_{r_1+1}^{(1)}; \gamma_1) \\
&\qquad = \left\{h_1(\boldsymbol{X}_1^{(1)}, \boldsymbol{X}_2^{(1)}; \boldsymbol{X}_3^{(1)}, \ldots, \boldsymbol{X}_{r_1+1}^{(1)}; \gamma_1) - p_1(\boldsymbol{X}_1^{(1)}, \boldsymbol{X}_2^{(1)}; \gamma_1)\right\} \\
&\qquad\quad \times \sin(\pi p_1(\boldsymbol{X}_1^{(1)}, \boldsymbol{X}_2^{(1)}; \gamma_1)),
\end{aligned}$$

and g_2^* is a symmetrized version of the kernel g_2. $U_{2n}(\boldsymbol{\theta}_1)$ is at least a first-order degenerate U-statistic, which is seen by observing that (taking $\boldsymbol{\theta}_1 = \mathbf{0}$)

$$\begin{aligned}
\mathrm{E}_{H_0}\left[g_2^*(\boldsymbol{X}_1^{(1)},\dots,\boldsymbol{X}_{r_1+1}^{(1)}) \mid \boldsymbol{X}_1^{(1)}\right] &= a_1 \mathrm{E}_{H_0}\left[g_2(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)};\boldsymbol{X}_3^{(1)},\dots,\boldsymbol{X}_{r_1+1}^{(1)}) \mid \boldsymbol{X}_1^{(1)}\right] \\
&\quad + a_2 \mathrm{E}_{H_0}\left[g_2(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)};\boldsymbol{X}_3^{(1)},\dots,\boldsymbol{X}_{r_1+1}^{(1)}) \mid \boldsymbol{X}_3^{(1)}\right] \\
&= a_1 \cdot 0 + a_2 \cdot 0 = 0.
\end{aligned}$$

The latter follows from the fact that

$$\begin{aligned}
&\mathrm{E}_{H_0}\left[g_2(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)};\boldsymbol{X}_3^{(1)},\dots,\boldsymbol{X}_{r_1+1}^{(1)}) \mid \boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)}\right] \\
&= \mathrm{E}_{H_0}\Big[\left\{h_1(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)};\boldsymbol{X}_3^{(1)},\dots,\boldsymbol{X}_{r_1+1}^{(1)}) - p_1(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)})\right\} \\
&\qquad \times \sin(\pi p_1(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)})) \mid \boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)}\Big] \\
&= \left\{\mathrm{E}_{H_0}\left[h_1(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)};\boldsymbol{X}_3^{(1)},\dots,\boldsymbol{X}_{r_1+1}^{(1)}) \mid \boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)}\right] - p_1(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)})\right\} \\
&\qquad \times \sin(\pi p_1(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)})) \\
&= 0 \cdot \sin(\pi p_1(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)})) \qquad \text{(by definition)} \\
&= 0,
\end{aligned}$$

and since the $\boldsymbol{X}^{(1)}$'s are jointly symmetric about the origin,

$$
\begin{aligned}
&\mathrm{E}_{H_0}\left[g_2(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)};\boldsymbol{X}_3^{(1)},\ldots,\boldsymbol{X}_{r_1+1}^{(1)}) \mid \boldsymbol{X}_3^{(1)},\ldots,\boldsymbol{X}_{r_1+1}^{(1)}\right] \\
&= \mathrm{E}_{H_0}\Big[\left\{h_1(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)};\boldsymbol{X}_3^{(1)},\ldots,\boldsymbol{X}_{r_1+1}^{(1)}) - p_1(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)})\right\} \\
&\qquad \times \sin(\pi p_1(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)})) \mid \boldsymbol{X}_3^{(1)},\ldots,\boldsymbol{X}_{r_1+1}^{(1)}\Big] \\
&= \mathrm{E}_{H_0}\Big[\left\{h_1(-\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)};\boldsymbol{X}_3^{(1)},\ldots,\boldsymbol{X}_{r_1+1}^{(1)}) - p_1(-\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)})\right\} \\
&\qquad \times \sin(\pi p_1(-\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)})) \mid \boldsymbol{X}_3^{(1)},\ldots,\boldsymbol{X}_{r_1+1}^{(1)}\Big] \\
&= \mathrm{E}_{H_0}\Big[\left\{(1 - h_1(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)};\boldsymbol{X}_3^{(1)},\ldots,\boldsymbol{X}_{r_1+1}^{(1)})) - (1 - p_1(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)}))\right\} \\
&\qquad \times \sin(\pi(1 - p_1(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)}))) \mid \boldsymbol{X}_3^{(1)},\ldots,\boldsymbol{X}_{r_1+1}^{(1)}\Big] \\
&= -\mathrm{E}_{H_0}\Big[\left\{h_1(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)};\boldsymbol{X}_3^{(1)},\ldots,\boldsymbol{X}_{r_1+1}^{(1)})) - p_1(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)})\right\} \\
&\qquad \times \sin(\pi p_1(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)})) \mid \boldsymbol{X}_3^{(1)},\ldots,\boldsymbol{X}_{r_1+1}^{(1)}\Big] \\
&= -\mathrm{E}_{H_0}\left[g_2(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)};\boldsymbol{X}_3^{(1)},\ldots,\boldsymbol{X}_{r_1+1}^{(1)}) \mid \boldsymbol{X}_3^{(1)},\ldots,\boldsymbol{X}_{r_1+1}^{(1)}\right],
\end{aligned}
$$

which implies that $\mathrm{E}_{H_0}\left[g_2(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)};\boldsymbol{X}_3^{(1)},\ldots,\boldsymbol{X}_{r_1+1}^{(1)}) \mid \boldsymbol{X}_3^{(1)},\ldots,\boldsymbol{X}_{r_1+1}^{(1)}\right] = 0$. Then (Lee, 1990, p. 83) implies that $U_{2n}(\boldsymbol{\theta}_1) = O_p(n^{-1})$. Applying Theorem A.0.1 we have

$$
\begin{aligned}
U_{2n}(\hat{\boldsymbol{\theta}}_1) &= U_{2n}(\boldsymbol{\theta}_1) + o_p(n^{-1/2}) \\
&= O_p(n^{-1}) + o_p(n^{-1/2}) \\
&= o_p(n^{-1/2}),
\end{aligned}
$$

since $\mathrm{E}_{H_0}\left[g_2(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)};\boldsymbol{X}_3^{(1)},\ldots,\boldsymbol{X}_{r_1+1}^{(1)};\gamma_1) \mid \boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)}\right] = 0$ regardless of the value of γ_1, and hence $\mathrm{E}_{H_0}\left[U_{2n}(\gamma_1)\right] = 0$ for all γ_1. Thus $V_{2n}(\hat{\boldsymbol{\theta}}_1) = o_p(n^{-1/2})$. Condition 2 is satisfied if it holds for the kernel $h_1(\boldsymbol{x}_1^{(1)},\boldsymbol{x}_2^{(1)};\boldsymbol{x}_3^{(1)},\ldots,\boldsymbol{x}_{r_1+1}^{(1)};\gamma_1)$, which we label as condition (C3). We also need (C2) to hold again.

It turns out that $V_{3n}(\gamma_1)$ is bounded by a multiple of a U-statistic plus a positive constant.

$$V_{3n}(\gamma_1)$$

$$\leq C_n \binom{2r_1}{2}^{-1} \binom{2r_1-2}{r_1-1}^{-1} \binom{n}{2r_1}^{-1} \sum_{\beta \in C_n^{2r_1}} g_3^*(\boldsymbol{X}_{\beta_1}^{(1)}, \ldots, \boldsymbol{X}_{\beta_{2r_1}}^{(1)}; \gamma_1) + (1 - C_n)$$

$$= C_n \binom{2r_1-2}{r_1-1}^{-1} \binom{2r_1}{2}^{-1} U_3(\gamma_1) + (1 - C_n),$$

where

$$g_3(\boldsymbol{X}_1^{(1)}, \boldsymbol{X}_2^{(1)}; \boldsymbol{X}_3^{(1)}, \ldots, \boldsymbol{X}_{r_1+1}^{(1)}; \boldsymbol{X}_{r_1+2}^{(1)}, \ldots, \boldsymbol{X}_{2r_1}^{(1)}; \gamma_1)$$

$$= \left\{ h_1(\boldsymbol{X}_1^{(1)}, \boldsymbol{X}_2^{(1)}; \boldsymbol{X}_3^{(1)}, \ldots, \boldsymbol{X}_{r_1+1}^{(1)}; \gamma_1) - p_1(\boldsymbol{X}_1^{(1)}, \boldsymbol{X}_2^{(1)}; \gamma_1) \right\}$$

$$\times \left\{ h_1(\boldsymbol{X}_1^{(1)}, \boldsymbol{X}_2^{(1)}; \boldsymbol{X}_{r_1+2}^{(1)}, \ldots, \boldsymbol{X}_{2r_1}^{(1)}; \gamma_1) - p_1(\boldsymbol{X}_1^{(1)}, \boldsymbol{X}_2^{(1)}; \gamma_1) \right\},$$

with $C_n = \binom{n-r_1-1}{r_1-1} / \binom{n-2}{r_1-1}$ and g_3^* a symmetrized version of the kernel g_3. Some simple algebra reveals that $1 - C_n = O(n^{-1})$. Also, $U_{3n}(\boldsymbol{\theta}_1)$ is at least first-order degenerate following the same reasoning as for $U_{2n}(\boldsymbol{\theta}_1)$. We have (taking $\boldsymbol{\theta}_1 = \boldsymbol{0}$) that

$$\mathrm{E}_{H_0} \left[g_3^*(\boldsymbol{X}_1^{(1)}, \ldots, \boldsymbol{X}_{2r_1}^{(1)}) \mid \boldsymbol{X}_1^{(1)} \right]$$

$$= a_1 \mathrm{E}_{H_0} \left[g_3(\boldsymbol{X}_1^{(1)}, \boldsymbol{X}_2^{(1)}; \boldsymbol{X}_3^{(1)}, \ldots, \boldsymbol{X}_{r_1+1}^{(1)}; \boldsymbol{X}_{r_1+2}^{(1)}, \ldots, \boldsymbol{X}_{2r_1}^{(1)}) \mid \boldsymbol{X}_1^{(1)} \right]$$

$$+ a_2 \mathrm{E}_{H_0} \left[g_3(\boldsymbol{X}_1^{(1)}, \boldsymbol{X}_2^{(1)}; \boldsymbol{X}_3^{(1)}, \ldots, \boldsymbol{X}_{r_1+1}^{(1)}; \boldsymbol{X}_{r_1+2}^{(1)}, \ldots, \boldsymbol{X}_{2r_1}^{(1)}) \mid \boldsymbol{X}_3^{(1)} \right]$$

$$= a_1 \cdot 0 + a_2 \cdot 0 = 0,$$

which follows since

$$\mathrm{E}_{H_0}\left[g_3(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)};\boldsymbol{X}_3^{(1)},\ldots,\boldsymbol{X}_{r_1+1}^{(1)};\boldsymbol{X}_{r_1+2}^{(1)},\ldots,\boldsymbol{X}_{2r_1}^{(1)}) \mid \boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)},\boldsymbol{X}_3^{(1)},\ldots,\boldsymbol{X}_{r_1+1}^{(1)}\right]$$

$$= \mathrm{E}_{H_0}\Big[\left\{h_1(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)};\boldsymbol{X}_3^{(1)},\ldots,\boldsymbol{X}_{r_1+1}^{(1)}) - p_1(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)})\right\}$$

$$\times\left\{h_1(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)};\boldsymbol{X}_{r_1+2}^{(1)},\ldots,\boldsymbol{X}_{2r_1}^{(1)}) - p_1(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)})\right\}$$

$$\mid \boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)},\boldsymbol{X}_3^{(1)},\ldots,\boldsymbol{X}_{r_1+1}^{(1)}\Big]$$

$$= \left\{h_1(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)};\boldsymbol{X}_3^{(1)},\ldots,\boldsymbol{X}_{r_1+1}^{(1)}) - p_1(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)})\right\}$$

$$\times\left\{\mathrm{E}_{H_0}\left[h_1(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)};\boldsymbol{X}_{r_1+2}^{(1)},\ldots,\boldsymbol{X}_{2r_1}^{(1)}) \mid \boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)}\right] - p_1(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)})\right\}$$

$$= \left\{h_1(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)};\boldsymbol{X}_3^{(1)},\ldots,\boldsymbol{X}_{r_1+1}^{(1)}) - p_1(\boldsymbol{X}_1^{(1)},\boldsymbol{X}_2^{(1)})\right\}\cdot 0 \qquad \text{(by definition)}$$

$$= 0.$$

Again, (Lee, 1990, p. 83) shows that $U_{3n}(\boldsymbol{\theta}_1)$ is $O_p(n^{-1})$. Since $\mathrm{E}_{H_0}\left[U_{3n}(\boldsymbol{\gamma}_1)\right] = 0$ for all $\boldsymbol{\gamma}_1$, applying Theorem A.0.1 shows that

$$\begin{aligned} U_{3n}(\hat{\boldsymbol{\theta}}_1) &= U_{3n}(\boldsymbol{\theta}_1) + o_p(n^{-1/2}) \\ &= O_p(n^{-1}) + o_p(n^{-1/2}) \\ &= o_p(n^{-1/2}), \end{aligned}$$

so that

$$\begin{aligned} V_{3n}(\hat{\boldsymbol{\theta}}_1) &= (1 + O(n^{-1}))o_p(n^{-1/2}) + O(n^{-1}) \\ &= o_p(n^{-1/2}). \end{aligned}$$

No further conditions than those already stated for $V_{1n}(\hat{\boldsymbol{\theta}}_1)$ and $V_{2n}(\hat{\boldsymbol{\theta}}_1)$ are needed. Since $V_{1n}(\hat{\boldsymbol{\theta}}_1)$, $V_{2n}(\hat{\boldsymbol{\theta}}_1)$, and $V_{3n}(\hat{\boldsymbol{\theta}}_1)$ have all been shown to be $o_p(n^{-1/2})$, referring to (A.2) leads to the conclusion that $\hat{V}_n(\hat{\boldsymbol{\theta}}_1) = o_p(n^{-1/2})$ as well. □

Lemma A.0.2 $\bar{d} = o_p(n^{-1/2})$.

Proof of Lemma A.0.2 After some algebra, the proof is quite similar to the proof of Lemma A.0.1, (with similar assumptions) so it is omitted. □

Lemma A.0.3 $c_2 = o_p(n^{5/2})$.

Proof of Lemma A.0.3 In light of the limiting behaviour of $\bar{c}$, we need only work with

$$\sum_{\boldsymbol{\beta} \in D_n^3} c(\beta_1, \beta_2) c(\beta_1, \beta_3).$$

Define

$$\hat{W}_n(\gamma_1) = \binom{n}{3}^{-1} \sum_{\boldsymbol{\beta} \in D_n^3} \cos(\pi \hat{p}_1(\beta_1, \beta_2; \gamma_1)) \cos(\pi \hat{p}_1(\beta_1, \beta_3; \gamma_1)).$$

Again using a Taylor series expansion of $\cos x$ and previous limiting results, it suffices to show that each of $W_{1n}(\hat{\boldsymbol{\theta}}_1)$, $W_{2n}(\hat{\boldsymbol{\theta}}_1)$, and $W_{3n}(\hat{\boldsymbol{\theta}}_1)$ is $o_p(n^{-1/2})$, where

$$W_{1n}(\gamma_1) = \binom{n}{3}^{-1} \sum_{\boldsymbol{\beta} \in D_n^3} \cos(\pi p_1(\beta_1, \beta_2; \gamma_1)) \cos(\pi p_1(\beta_1, \beta_3; \gamma_1)),$$

$$W_{2n}(\gamma_1) = \binom{n}{3}^{-1} \sum_{\boldsymbol{\beta} \in D_n^3} \left\{\hat{p}_1(\beta_1, \beta_2; \gamma_1) - p_1(\beta_1, \beta_2; \gamma_1)\right\}$$

$$\times \sin(\pi p_1(\beta_1, \beta_2; \gamma_1)) \cos(\pi p_1(\beta_1, \beta_3; \gamma_1)),$$

and

$$W_{3n}(\gamma_1) = \binom{n}{3}^{-1} \sum_{\boldsymbol{\beta} \in D_n^3} \left\{\hat{p}_1(\beta_1, \beta_2; \gamma_1) - p_1(\beta_1, \beta_2; \gamma_1)\right\}$$

$$\times \left\{\hat{p}_1(\beta_1, \beta_3; \gamma_1) - p_1(\beta_1, \beta_3; \gamma_1)\right\}$$

$$\times \sin(\pi p_1(\beta_1, \beta_2; \gamma_1)) \sin(\pi p_1(\beta_1, \beta_3; \gamma_1)).$$

Analysis similar to that in the proof of Lemma A.0.1 using the repeated application of Theorem A.0.1 (and with no further assumptions than those needed in Lemma A.0.1) proves the result. □

Lemma A.0.4 $d_2 = o_p(n^{5/2})$.

Proof of Lemma A.0.4 After some algebra, the proof is again quite similar to the proof of Lemma A.0.3, so it also is omitted. □

Lemma A.0.5 $c_1 = O(n^2)$.

Proof of Lemma A.0.5 It suffices to note that $c(i,j)$ and $\bar{c}$ are both bounded. □

In order to simplify the proof of $d_1 = o_p(n^2)$, and because it will be of general use, we prove the following lemma.

Lemma A.0.6 $\hat{p}_k(i,j;\hat{\boldsymbol{\theta}}_k) = \hat{p}_k(i,j;\boldsymbol{\theta}_k) + o_p(1)$.

Proof of Lemma A.0.6 Let $\epsilon > 0$, $\delta > 0$, and Condition 2 be satisfied with h replaced by $h_k(\boldsymbol{x}_1^{(k)}, \boldsymbol{x}_2^{(k)}; \boldsymbol{x}_3^{(k)}, \ldots, \boldsymbol{x}_{\tau_k+1}^{(k)}; \gamma_k)$. Then we can choose a sphere $D(\boldsymbol{\theta}_k, d^*)$ so that

$$B \equiv \mathrm{E}_{H_0}\left[\sup_{\gamma_k \in D(\boldsymbol{\theta}_k, d^*)} \left| h_k(\boldsymbol{X}_1^{(k)}, \boldsymbol{X}_2^{(k)}; \boldsymbol{X}_3^{(k)}, \ldots, \boldsymbol{X}_{\tau_k+1}^{(k)}; \gamma_k) \right.\right.$$
$$\left.\left. - h_k(\boldsymbol{X}_1^{(k)}, \boldsymbol{X}_2^{(k)}; \boldsymbol{X}_3^{(k)}, \ldots, \boldsymbol{X}_{\tau_k+1}^{(k)}; \boldsymbol{\theta}_k) \right|\right] \leq \epsilon/2.$$

Now,

$$\mathrm{P}\left[|\hat{p}_k(i,j;\hat{\boldsymbol{\theta}}_k) - \hat{p}_k(i,j;\boldsymbol{\theta}_k)| > \epsilon\right] \leq \mathrm{P}\left[\sup_{||\mathbf{t}|| \leq d^*} |\hat{p}_k(i,j;\boldsymbol{\theta}_k + \mathbf{t}) - \hat{p}_k(i,j;\boldsymbol{\theta}_k)| > \epsilon\right]$$
$$+ \mathrm{P}\left[||\hat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}_k|| > d^*\right]$$

which for n sufficiently large,

$$\leq \mathrm{P}\left[\sup_{||\mathbf{t}|| \leq d^*} |\hat{p}_k(i,j;\boldsymbol{\theta}_k + \mathbf{t}) - \hat{p}_k(i,j;\boldsymbol{\theta}_k)| > \epsilon\right] + \delta/2$$
$$= \mathrm{P}\left[U_n^* > \epsilon\right] + \delta/2,$$

where U_n^* is almost a U-statistic ($\boldsymbol{X}_1^{(k)}$ and $\boldsymbol{X}_2^{(k)}$ are common to all terms) which estimates B. Using (Serfling, 1980, Theorem A, p. 201) yields

$$\begin{aligned}
\mathrm{P}\left[U_n^* - B > \epsilon/2\right] &= \mathrm{E}_{H_0}\left[\mathrm{E}_{H_0}\left[I(U_n^* - B > \epsilon/2) \mid \boldsymbol{X}_i^{(k)}, \boldsymbol{X}_j^{(k)}\right]\right] \\
&= \mathrm{E}_{H_0}\left[\mathrm{P}\left[U_n^* - B > \epsilon/2 \mid \boldsymbol{X}_i^{(k)}, \boldsymbol{X}_j^{(k)}\right]\right] \\
&\leq \mathrm{E}_{H_0}\left[\exp(-(1/2)\lfloor n/(r_2-1)\rfloor(\epsilon/2)^2)\right] \\
&= \exp(-\lfloor n/(r_2-1)\rfloor(\epsilon^2/8))
\end{aligned}$$

which for n sufficiently large,

$$< \delta/2.$$

Thus for n sufficiently large,

$$\begin{aligned}
\mathrm{P}\left[|\hat{p}_k(i,j;\hat{\boldsymbol{\theta}}_k) - \hat{p}_k(i,j;\boldsymbol{\theta}_k)| > \epsilon\right] &\leq \mathrm{P}\left[U_n^* > \epsilon\right] + \delta/2 \\
&\leq \mathrm{P}\left[U_n^* - B > \epsilon/2\right] + \delta/2 \qquad (\text{since } \epsilon/2 + B \leq \epsilon) \\
&< \delta/2 + \delta/2 = \delta. \qquad \square
\end{aligned}$$

<u>Lemma A.0.7</u> $d_1 = o_p(n^2)$.

<u>Proof of Lemma A.0.7</u> Since $\bar{d} = o_p(n^{-1/2})$ and $d(i,j) \leq 2$, it suffices to show that $\sum_{i \neq j}^{n} |d(i,j)| = o_p(n^2)$. This will follow if $\mathrm{E}_{H_0}\left[|d(i,j)|\right] \to 0$ as $n \to \infty$ because $d(i,j)$ is bounded. Now

$$\begin{aligned}
\mathrm{E}_{H_0}\left[|d(i,j)|\right] &= \mathrm{E}_{H_0}\left[|\cos(\pi\hat{p}_2(i,j;\hat{\boldsymbol{\theta}}_2)) - \cos(\pi\hat{p}_2(i,j;\boldsymbol{\theta}_2))|\right] \\
&\leq \pi\mathrm{E}_{H_0}\left[|\hat{p}_2(i,j;\hat{\boldsymbol{\theta}}_2) - \hat{p}_2(i,j;\boldsymbol{\theta}_2)|\right] \\
&\to 0,
\end{aligned}$$

with the last statement following from Lemma A.0.6 and the fact that the integrand is bounded. $\square$

We now sketch an argument that in the bivariate case, (C1), (C2) and (C3) are attainable conditions with some mild assumptions on the underlying density f. Let D and D' be circles with radii d and $2d$, respectively, centered at γ, and contained within the circle K. Restrict K to be inside the unit circle. Let $A_1(\boldsymbol{x}, D)$ be equal to the probability that a random variable $\boldsymbol{Z}$ with absolutely continuous density function f falls in one of the two extended triangular regions above and below $\boldsymbol{x}$ as shown in Figure A.1. Then for a given $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$,

$$\sup_{\gamma' \in D} |p(\boldsymbol{x}_1, \boldsymbol{x}_2; \gamma) - p(\boldsymbol{x}_1, \boldsymbol{x}_2; \gamma')| \leq A_1(\boldsymbol{x}_1, D) I(\boldsymbol{x}_1 \notin D') + A_1(\boldsymbol{x}_2, D) I(\boldsymbol{x}_2 \notin D')$$
$$+ I(\boldsymbol{x}_1 \text{ or } \boldsymbol{x}_2 \in D'),$$

so that it suffices to show $\mathrm{E}_{H_0}[A_1(\boldsymbol{X}, D) I(\boldsymbol{X} \notin D')] \leq k_1 d$ and $P(\boldsymbol{X} \in D') \leq k_2 d$ for some constants k_1 and k_2 in order to satisfy (C2), where $\boldsymbol{X}$ also has the density f. Assume that f is spherically symmetric so that $f(\boldsymbol{x}) \propto g(r^2)$, where $r = ||\boldsymbol{x}||$. Further assume that $g(r^2)$ is non-increasing in r^2 (unimodal) with $g(0)$ finite (so that f is bounded). Then $A_1(\boldsymbol{x}, D)$ is maximized when γ is on the line through $\boldsymbol{x}$ and the origin. Referring to Figure A.2, $A_1(\boldsymbol{x}, D) < 4A_2(\boldsymbol{x}, D)$, where $A_2(\boldsymbol{x}, D)$ is the probability that $\boldsymbol{Z}$ falls in the extended triangular region below $\boldsymbol{x}$. Now for $\boldsymbol{x} \notin D'$, (i.e., $||\boldsymbol{x} - \gamma|| > 2d$),

$$\alpha = \sin^{-1}(||\boldsymbol{x} - \gamma||^{-1} d) \leq (2/\sqrt{3}) ||\boldsymbol{x} - \gamma||^{-1} d.$$

It then follows using some basic trigonometry (see Figure A.3) that

$$A_2(\boldsymbol{x}, D) \leq (4/\sqrt{3}\pi) ||\boldsymbol{x} - \gamma||^{-1} d + (2r^2/\sqrt{3}) ||\boldsymbol{x} - \gamma||^{-1} d.$$

Thus, we need $\mathrm{E}_{H_0}[||\boldsymbol{X} - \gamma||^{-1}] < \infty$ and $\mathrm{E}_{H_0}[R^2 ||\boldsymbol{X} - \gamma||^{-1}] < \infty$ ($R = ||\boldsymbol{X}||$). Both of these moments can be shown to be finite if in addition to the assumptions on f we add $\mathrm{E}_{H_0}[R^2] < \infty$. This can be seen as follows. Clearly $\mathrm{E}_{H_0}[||\boldsymbol{X} - \gamma||^{-1}] < \infty$ if and only if

$$\mathrm{E}_{H_0}\left[||\boldsymbol{X} - \gamma||^{-1} I(||\boldsymbol{X} - \gamma|| < c)\right] < \infty$$

for c a positive number. Because of the symmetry of the problem, without loss of generality take $\boldsymbol{\gamma} = (\gamma, 0)'$ and $0 < \gamma < 1/2$. For $c = 1/2$,

$$\begin{aligned}
&\mathrm{E}_{H_0}\left[||\boldsymbol{X} - \boldsymbol{\gamma}||^{-1} I(||\boldsymbol{X} - \boldsymbol{\gamma}|| < c)\right] \\
&\quad = \int\int ((x_1 - \gamma)^2 + x_2^2)^{-1/2} I(||\boldsymbol{X} - \boldsymbol{\gamma}|| < 1/2) f(x_1, x_2)\, dx_1 dx_2 \\
&\quad \leq f(0,0) \int\int ((x_1 - \gamma)^2 + x_2^2)^{-1/2} I(0 < x_1 < 1) I(0 < x_2 < 1)\, dx_1 dx_2 \\
&\quad = f(0,0) \int_{-1}^{1} \int_{-1}^{1} ((x_1 - \gamma)^2 + x_2^2)^{-1/2}\, dx_2 dx_1.
\end{aligned}$$

This last integral can easily be verified finite, and again similar reasoning shows that $\mathrm{E}_{H_0}\,[R^2||\boldsymbol{X} - \boldsymbol{\gamma}||^{-1}] < \infty$. Further, since $P(\boldsymbol{X} \in D') \leq (4\pi f(0))^2 d^2$, we have $P(\boldsymbol{X} \in D') \leq k_2 d$ since $d < 1$. Similar arguments can be used to show (C3) holds as well.

To show (C1) can reasonably be attained, we utilize results pertaining to the continuity and differentiability of integrals, for example (Randles & Wolfe, 1979, Theorems A.2.3 and A.2.4, pp. 417–418). First observe that

$$\int h(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{z} - \boldsymbol{\gamma}) f(\boldsymbol{z})\, d\boldsymbol{z} = \int h(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{z}) f(\boldsymbol{z} + \boldsymbol{\gamma})\, d\boldsymbol{z},$$

where h is the indicator of whether $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are on opposite sides of the line formed by $\boldsymbol{z} - \boldsymbol{\gamma}$ and the origin. This integral is a continuous function of $\boldsymbol{\gamma}$ in a neighborhood of the origin, since we can bound the integrand by $f(\boldsymbol{z})$, with $\int f(\boldsymbol{z}) d\boldsymbol{z} = 1 < \infty$. Further, if the elements of $\nabla f(\boldsymbol{z} + \boldsymbol{\gamma})$ can be bounded by an integrable function not depending on $\boldsymbol{\gamma}$, then

$$\nabla \int h(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{z}) f(\boldsymbol{z} + \boldsymbol{\gamma})\, d\boldsymbol{z} = \int h(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{z}) \nabla f(\boldsymbol{z} + \boldsymbol{\gamma})\, d\boldsymbol{z}$$

will exist and be continuous. Now

$$\begin{aligned}
&\int\int \cos\left(\pi \int h(\boldsymbol{x}_1 - \boldsymbol{\gamma}, \boldsymbol{x}_2 - \boldsymbol{\gamma}, \boldsymbol{z} - \boldsymbol{\gamma}) f(\boldsymbol{z})\, d\boldsymbol{z}\right) f(\boldsymbol{x}_1) f(\boldsymbol{x}_2)\, d\boldsymbol{x}_1 d\boldsymbol{x}_2 = \\
&\qquad \int\int \cos\left(\pi \int h(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{z}) f(\boldsymbol{z} + \boldsymbol{\gamma})\, d\boldsymbol{z}\right) f(\boldsymbol{x}_1 + \boldsymbol{\gamma}) f(\boldsymbol{x}_2 + \boldsymbol{\gamma})\, d\boldsymbol{x}_1 d\boldsymbol{x}_2,
\end{aligned}$$

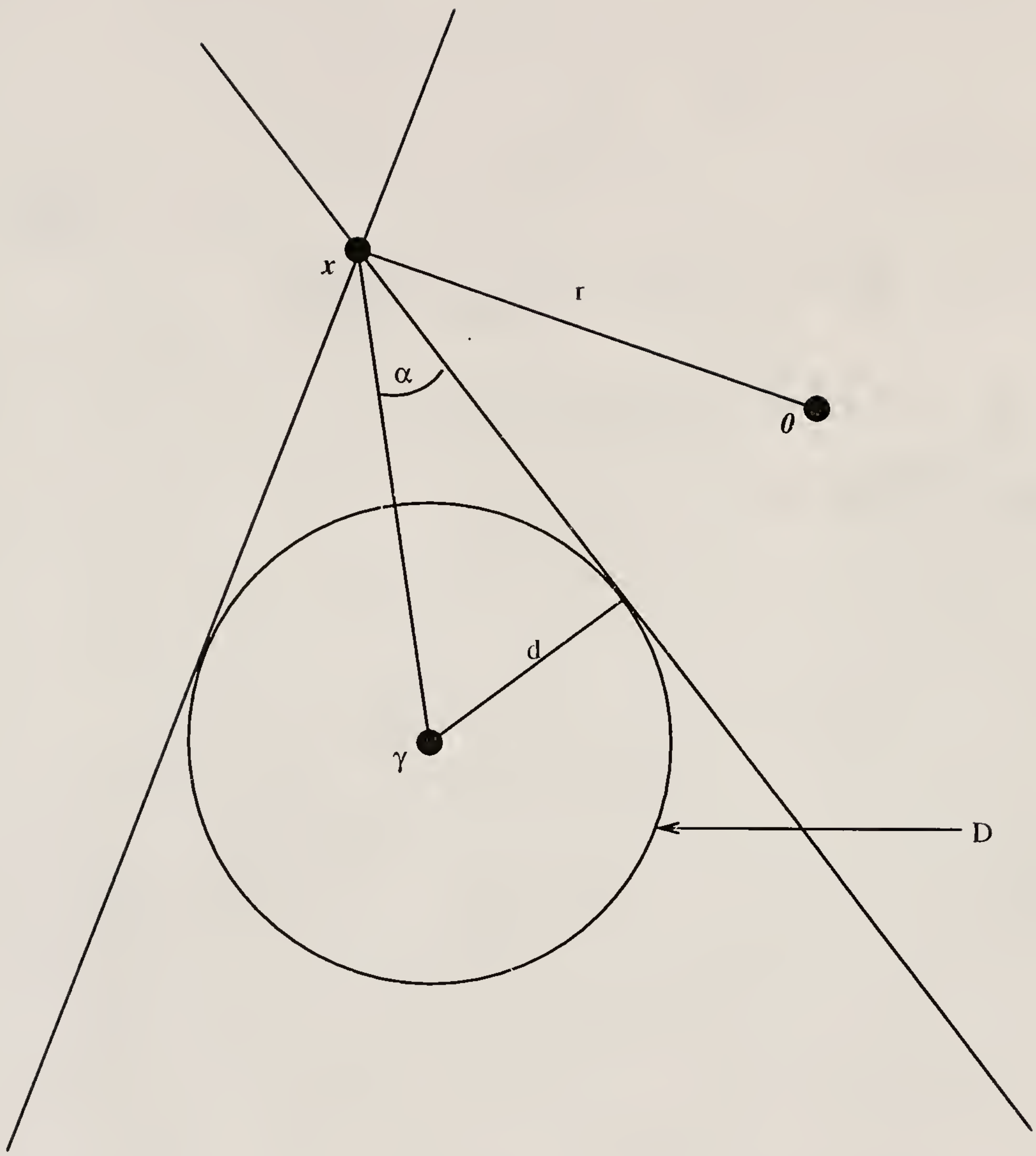

Figure A.1. Region 1

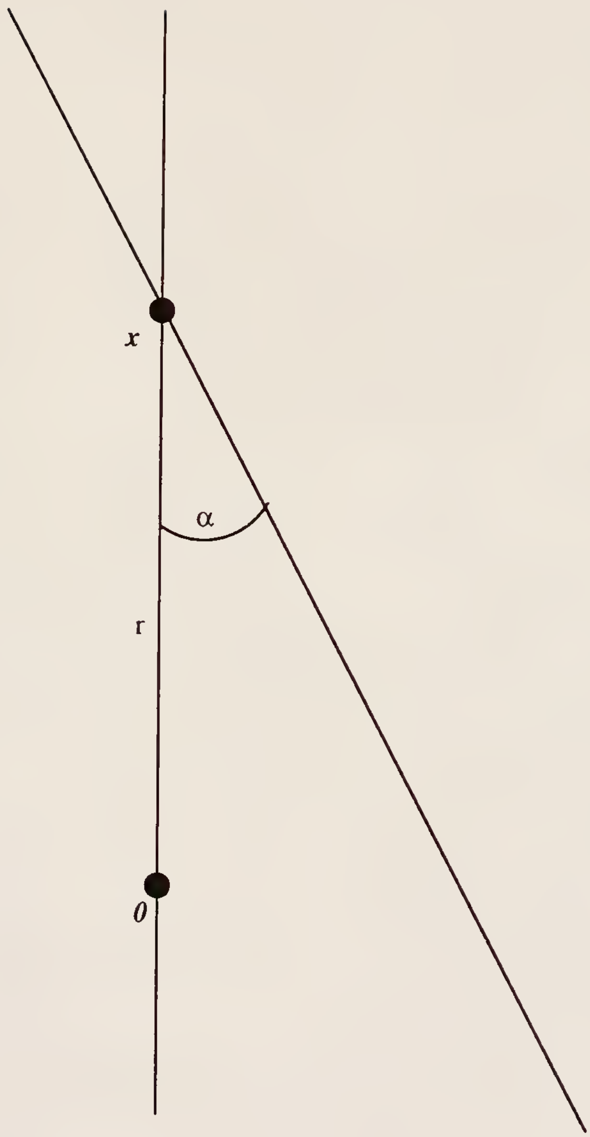

Figure A.2. Region 2

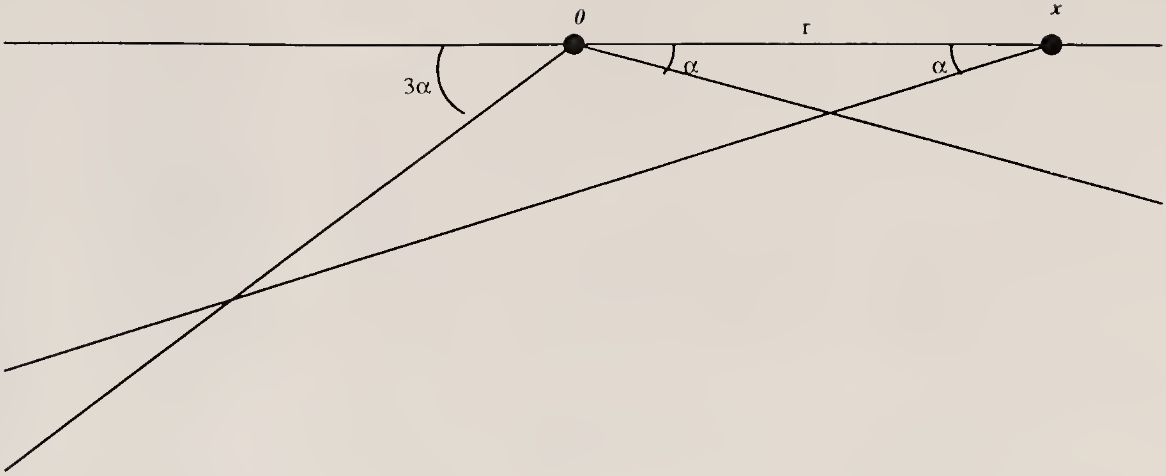

Figure A.3. Region 3

where the integrand is a continuous function of γ is the mean function $\mu_{V_{1n}}(\gamma)$. By similar reasoning, $\nabla \mu_{V_{1n}}(\gamma)$ is given by

$$\nabla \int\int \cos\left(\pi \int h(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{z}) f(\boldsymbol{z}+\gamma)\, d\boldsymbol{z}\right) f(\boldsymbol{x}_1+\gamma) f(\boldsymbol{x}_2+\gamma)\, d\boldsymbol{x}_1 d\boldsymbol{x}_2 =$$

$$\int\int \sin\left(\pi \int h(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{z}) f(\boldsymbol{z}+\gamma)\, d\boldsymbol{z}\right) \cdot \left(\pi \int h(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{z}) \nabla f(\boldsymbol{z}+\gamma)\, d\boldsymbol{z}\right)\, d\boldsymbol{x}_1 d\boldsymbol{x}_2$$

$$+2 \int\int \cos\left(\pi \int h(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{z}) f(\boldsymbol{z}+\gamma)\, d\boldsymbol{z}\right) \nabla f(\boldsymbol{x}_1+\gamma) f(\boldsymbol{x}_2+\gamma)\, d\boldsymbol{x}_1 d\boldsymbol{x}_2$$

will exist and be continuous if we can bound the components of $f(\boldsymbol{z}+\gamma)$ and $\nabla f(\boldsymbol{z}+\gamma)$ by integrable functions not depending on γ. These conditions on f are satisfied in the case of the normal distribution among others. For example, letting $f(x, y)$ denote the bivariate standard normal density, we can define the dominating function g as

follows:

$$g(x,y) = \begin{cases} f(0,0) & \begin{cases} -1 < x < 1 \\ -1 < y < 1 \end{cases} \\ f(0,y+1) & \begin{cases} -1 < x < 1 \\ -\infty < y < -1 \end{cases} \\ f(0,y-1) & \begin{cases} -1 < x < 1 \\ 1 < y < \infty \end{cases} \\ f(x+1,0) & \begin{cases} -\infty < x < -1 \\ -1 < y < 1 \end{cases} \\ f(x-1,0) & \begin{cases} 1 < x < \infty \\ -1 < y < 1 \end{cases} \\ f(x+1,y-1) & \begin{cases} -\infty < x < -1 \\ 1 < y < \infty \end{cases} \\ f(x+1,y+1) & \begin{cases} -\infty < x < -1 \\ -\infty < y < -1 \end{cases} \\ f(x-1,y-1) & \begin{cases} 1 < x < \infty \\ 1 < y < \infty \end{cases} \\ f(x-1,y+1) & \begin{cases} 1 < x < \infty \\ -\infty < y < -1 \end{cases} \end{cases}.$$

Clearly, this function dominates $f(x+\gamma_1, y+\gamma_2)$ for $||\gamma|| < 1$ and is integrable because the tail behavior is essentially the same as for the original normal random variable. Because $\nabla f(\boldsymbol{x}+\boldsymbol{\gamma}) = -(\boldsymbol{x}+\boldsymbol{\gamma})f(\boldsymbol{x}+\boldsymbol{\gamma})$, similar reasoning will lead to an integrable dominating function. Thus, we know the conditions used in the technical lemmas are achievable.

APPENDIX B
CONTIGUITY

We need to establish the two propositions

- $W_n = T_n - \sigma^2/4 + o_p(1)$
- The summands $L(\boldsymbol{X}_i; \Delta_n)$ are UAN (i.e. $L(\boldsymbol{X}_i; \Delta_n) = 1 + o_p(1)$)

when the null hypothesis is true and $\mathrm{V}_{H_0}[T_n] \equiv \sigma^2 < \infty$. Keep in mind that we have $\Delta_n = n^{-1/2}\Delta_0$, where Δ_0 is a positive constant. We start with the first proposition. It suffices to show $\mathrm{E}_{H_0}[W_n] \to -\sigma^2/4$ and $\mathrm{V}_{H_0}[W_n - T_n] \to 0$ as $n \to \infty$. Let $s(\boldsymbol{x}) = f^{1/2}(\boldsymbol{x})$, so that $\nabla s(\boldsymbol{x}) = (\nabla f(\boldsymbol{x}))/(2f^{1/2}(\boldsymbol{x}))$. Then we can express $\mathrm{E}_{H_0}[W_n]$ as

$$\begin{aligned}
\mathrm{E}_{H_0}[W_n] &= 2n\mathrm{E}_{H_0}\left[L^{1/2}(\boldsymbol{X}; \Delta_n)\right] - 2 \\
&= 2n\mathrm{E}_{H_0}\left[\frac{|\boldsymbol{A}_{\Delta_n}|^{-1/2} s(\boldsymbol{A}_{\Delta_n}^{-1}\boldsymbol{X})}{s(\boldsymbol{X})}\right] - 2 \\
&= -\Delta_0^2 \int \left[\frac{|\boldsymbol{A}_{\Delta_n}|^{-1/2} s(\boldsymbol{A}_{\Delta_n}^{-1}\boldsymbol{x}) - s(\boldsymbol{x})}{\Delta_n}\right]^2 d\boldsymbol{x}.
\end{aligned}$$

We need to show that the integral converges to $\sigma^2/(4\Delta_0^2)$. Since

$$\lim_{\Delta_n \to 0}\left[\frac{|\boldsymbol{A}_{\Delta_n}|^{-1/2} s(\boldsymbol{A}_{\Delta_n}^{-1}\boldsymbol{x}) - s(\boldsymbol{x})}{\Delta_n}\right] = (1/2)(r_1 + r_2)s(\boldsymbol{x}) + (\boldsymbol{P}\boldsymbol{x})'\nabla s(\boldsymbol{x}),$$

and

$$\left[\frac{|\boldsymbol{A}_{\Delta_n}|^{-1/2}s(\boldsymbol{A}_{\Delta_n}^{-1}\boldsymbol{x}) - s(\boldsymbol{x})}{\Delta_n}\right]^2$$

$$= \left[\frac{1}{\Delta_n}\int_0^{\Delta_n}(1/2)|\boldsymbol{A}_t|^{-1/2}\mathrm{tr}(\boldsymbol{A}_t^{-1}\boldsymbol{P})s(\boldsymbol{x}) + |\boldsymbol{A}_t|^{-1/2}(\boldsymbol{A}_t^{-1}\boldsymbol{P}\boldsymbol{A}_t^{-1}\boldsymbol{x})'\nabla s(\boldsymbol{x})dt\right]^2$$

$$\leq \frac{1}{\Delta_n}\int_0^{\Delta_n}\left[(1/2)|\boldsymbol{A}_t|^{-1/2}\mathrm{tr}(\boldsymbol{A}_t^{-1}\boldsymbol{P})s(\boldsymbol{x}) + |\boldsymbol{A}_t|^{-1/2}(\boldsymbol{A}_t^{-1}\boldsymbol{P}\boldsymbol{A}_t^{-1}\boldsymbol{x})'\nabla s(\boldsymbol{x})\right]^2 dt,$$

we have

$$\int\left[\frac{|\boldsymbol{A}_{\Delta_n}|^{-1/2}s(\boldsymbol{A}_{\Delta_n}^{-1}\boldsymbol{x}) - s(\boldsymbol{x})}{\Delta_n}\right]^2 d\boldsymbol{x}$$

$$\leq \frac{1}{\Delta_n}\int\int_0^{\Delta_n}\left[(1/2)|\boldsymbol{A}_t|^{-1/2}\mathrm{tr}(\boldsymbol{A}_t^{-1}\boldsymbol{P})s(\boldsymbol{x}) + |\boldsymbol{A}_t|^{-1/2}(\boldsymbol{A}_t^{-1}\boldsymbol{P}\boldsymbol{A}_t^{-1}\boldsymbol{x})'\nabla s(\boldsymbol{x})\right]^2 dt\, d\boldsymbol{x}$$

$$\leq \frac{1}{\Delta_n}\int_0^{\Delta_n}\int\left[(1/2)|\boldsymbol{A}_t|^{-1/2}\mathrm{tr}(\boldsymbol{A}_t^{-1}\boldsymbol{P})s(\boldsymbol{x}) + |\boldsymbol{A}_t|^{-1/2}(\boldsymbol{A}_t^{-1}\boldsymbol{P}\boldsymbol{A}_t^{-1}\boldsymbol{x})'\nabla s(\boldsymbol{x})\right]^2 d\boldsymbol{x}\, dt$$

$$= \int\left[(1/2)(r_1 + r_2)s(\boldsymbol{x}) + (\boldsymbol{P}\boldsymbol{x})'\nabla s(\boldsymbol{x})\right]^2 d\boldsymbol{x},$$

so that

$$\lim_{n\to\infty}\mathrm{E}_{H_0}[W_n] = -\Delta_0^2\int\left[(1/2)(r_1 + r_2)s(\boldsymbol{x}) + (\boldsymbol{P}\boldsymbol{x})'\nabla s(\boldsymbol{x})\right]^2 d\boldsymbol{x}$$

$$= -\frac{\Delta_0^2}{4}\mathrm{E}_{H_0}\left[(L'(\boldsymbol{X};0))^2\right].$$

We need to show $\mathrm{E}_{H_0}[(L'(\boldsymbol{X};0))^2] = \sigma^2/\Delta_0^2$. This is clear, since

$$\mathrm{E}_{H_0}[L'(\boldsymbol{X};0)] = \int(r_1 + r_2)f(\boldsymbol{x}) + (\boldsymbol{P}\boldsymbol{x})'\nabla f(\boldsymbol{x})d\boldsymbol{x}$$

$$= \int\lim_{\Delta\to 0}\left[\frac{|\boldsymbol{A}_\Delta|^{-1}f(\boldsymbol{A}_\Delta^{-1}\boldsymbol{x}) - f(\boldsymbol{x})}{\Delta}\right]d\boldsymbol{x}$$

$$= \lim_{\Delta\to 0}(1/\Delta)\int[|\boldsymbol{A}_\Delta|^{-1}f(\boldsymbol{A}_\Delta^{-1}\boldsymbol{x}) - f(\boldsymbol{x})]d\boldsymbol{x}$$

$$= \lim_{\Delta\to 0}(1/\Delta)(1 - 1)$$

$$= 0,$$

implies that $\mathrm{E}_{H_0}[T_n] = 0$, and hence that $\sigma^2 = \Delta_0^2 \mathrm{E}_{H_0}[(L'(\boldsymbol{X};0))^2]$. Now since

$$\begin{aligned}
&\mathrm{V}_{H_0}[W_n - T_n] \\
&\quad = 4\Delta_0^2 \mathrm{V}_{H_0}\left[\frac{|\boldsymbol{A}_{\Delta_n}|^{-1/2} s(\boldsymbol{A}_{\Delta_n}^{-1}\boldsymbol{X})}{\Delta_n s(\boldsymbol{X})} - 1 + \frac{r_1 + r_2}{2} + \frac{(\boldsymbol{P}\boldsymbol{X})'\nabla s(\boldsymbol{X})}{s(\boldsymbol{X})}\right] \\
&\quad \leq 4\Delta_0^2 \mathrm{E}_{H_0}\left[\left(\frac{|\boldsymbol{A}_{\Delta_n}|^{-1/2} s(\boldsymbol{A}_{\Delta_n}^{-1}\boldsymbol{X})}{\Delta_n s(\boldsymbol{X})} - 1 + \frac{r_1 + r_2}{2} + \frac{(\boldsymbol{P}\boldsymbol{X})'\nabla s(\boldsymbol{X})}{s(\boldsymbol{X})}\right)^2\right] \\
&\quad = 4\Delta_0^2 \int \left[\frac{|\boldsymbol{A}_{\Delta_n}|^{-1/2} s(\boldsymbol{A}_{\Delta_n}^{-1}\boldsymbol{x}) - s(\boldsymbol{x})}{\Delta_n} + \frac{(r_1 + r_2) s(\boldsymbol{x})}{2} + (\boldsymbol{P}\boldsymbol{x})'\nabla s(\boldsymbol{x})\right]^2 d\boldsymbol{x},
\end{aligned}$$

$\mathrm{V}_{H_0}[W_n - V_n] \to 0$ as $n \to \infty$. To show the second proposition, first define $Z = L'(\boldsymbol{X};0)$ and

$$Z_n = \frac{|\boldsymbol{A}_{\Delta_n}|^{-1}|f(\boldsymbol{A}_{\Delta_n}^{-1}\boldsymbol{x}) - f(\boldsymbol{x})}{\Delta_n f(\boldsymbol{x})}.$$

Since

$$\mathrm{E}_{H_0}\left[(Z_n - Z)^2\right] = \int \frac{1}{f(\boldsymbol{x})}\left[\frac{|\boldsymbol{A}_{\Delta_n}|^{-1}|f(\boldsymbol{A}_{\Delta_n}^{-1}\boldsymbol{x}) - f(\boldsymbol{x})}{\Delta_n} - \frac{L'(\boldsymbol{x};0)}{f(\boldsymbol{x})}\right]^2 d\boldsymbol{x},$$

clearly $\mathrm{E}_{H_0}[(Z_n - Z)^2] \to 0$ as $n \to \infty$, thus $Z_n = O_p(1)$. Then $L(\boldsymbol{X};\Delta_n) = 1 + n^{-1/2}\Delta_0 Z_n = 1 + O_p(n^{-1/2}) = 1 + o_p(1)$ as needed.

APPENDIX C
SIMULATION STUDY

C.1 Source Code

Following are listings of some of the programs used to run the simulation study. They are not intended to be complete in the sense that they can stand alone. However, they do give an indication of the logic and the methods used.

C.1.1 interdir.c

```
#include <stdio.h>
#include <math.h>
#include "machine.h"
#include "general.h"

/*
 * Given an (n x p) data matrix A, InterdirectionCounts will compute
 * the interdirection counts, placing them in the upper triangular
 * (n x n) matrix COUNTS. As a by product, the return value is the
 * total number of p-1 dimensional hyperplanes through zero that
 * could be constructed from the n data points.
 */

int InterdirectionCounts(MAT *A, MAT *COUNTS)
{
    VEC *DELTA;
    double sum, temp;
    int *hyper, *base, *distinct, *subs;
    int pair[2];
    int first_hyper, first_pair;
    int n, p, n_hyper;
    int i, j;
    int Perpendicular(MAT *, int *, VEC *);
    double OppositeSide(MAT *, int *, VEC *);
```

```
n = A→m;
p = A→n;

if (p > 1) {

    DELTA = get_vec(p);
    hyper = (int *)malloc(p * sizeof(int));
    base = (int *)malloc(n * sizeof(int));
    distinct = (int *)malloc(n * sizeof(int));
    subs = (int *)malloc(n * sizeof(int));

    /*
     * hyper[] is an array of dimension p-1 which contains
     * the labels of the points (row numbers of A) in the
     * current hyperplane.
     */

    n_hyper = 0;
    first_hyper = TRUE;
    while (NextSubset(hyper,n,p−1,first_hyper)) {
        first_hyper = FALSE;

        if (Perpendicular(A,hyper,DELTA)) {
            ++n_hyper;

            /*
             * distinct[] is an array of dimension n-(p-1) which
             * contains the labels of the points not in the current
             * hyperplane.
             */

            for (i = 1; i ≤ n; ++i) base[i−1] = TRUE;
            for (i = 1; i ≤ p−1; ++i) base[hyper[i−1]−1] = FALSE;
            for (i = 1, j = 1; i ≤ n; ++i) {
                if (base[i−1]) {
                    distinct[j−1] = i;
                    ++j;
                }
            }

            /*
             * pair[] is an array of dimension 2 which
             * contains 2 of the labels from distinct[]
             * This loop iterates over all possible
```

```
            * pairs.
            */

            first_pair = TRUE;
            while (NextSubset(subs,n-(p-1),2,first_pair)) {
               first_pair = FALSE;
               pair[0] = distinct[subs[0]-1];
               pair[1] = distinct[subs[1]-1];
               temp = OppositeSide(A,pair,DELTA);
               sum = temp + m_entry(COUNTS,pair[0]-1,pair[1]-1);
               m_set(COUNTS,pair[0]-1,pair[1]-1,sum);
            }
         }
      }
      freevec(DELTA);
      free((void *)hyper);
      free((void *)base);
      free((void *)distinct);
      free((void *)subs);
   }
   else {
      n_hyper = 1;
      for (i = 1; i <= n - 1; ++i) {
         for (j = i + 1; j <= n; ++j) {
            temp = -(double)(sign(m_entry(A,i-1,0))
               *sign(m_entry(A,j-1,0)) - 1)/2;
            m_set(COUNTS,i-1,j-1,temp);
         }
      }
   }

   return n_hyper;
}


/*
 * Computes a perpendicular vector (DELTA) to a hyperplane
 * defined by the rows of A indicated by the row labels in
 * hyperplane and the origin.
 * Returns 1 if hyperplane is well defined and 0 if points
 * are collinear.
 */

int Perpendicular(MAT *A, int *hyper, VEC *DELTA)
{
```

```
    MAT *X, *Y;
    VEC *DELTASTAR;
    VEC *TEMP1, *TEMP2;
    int i, p, ret_val;

    p = A→n;

    X = get_mat(p−1,p);
    Y = get_mat(p−1,p−1);
    DELTASTAR = get_vec(p−1);
    TEMP1 = get_vec(p−1);
    TEMP2 = get_vec(p);

    for (i = 1; i ≤ p − 1; ++i) {
        get_row(A,hyper[i−1]−1,TEMP2);
        set_row(X,i−1,TEMP2);
    }
    sub_matrix(X,0,1,p−2,p−1,Y);

    ret_val = 0;

    if (fabs(det(Y)) > MACHEPS) {
        m_inverse(Y,Y);
        get_col(X,0,TEMP1);
        mv_mlt(Y,TEMP1,DELTASTAR);
        for (i = 1; i ≤ p − 1; ++i) {
            v_set(DELTA,i,v_entry(DELTASTAR,i−1));
        }
        v_set(DELTA,0,−1);
        ret_val = 1;
    }

    freemat(X);
    freemat(Y);
    freevec(DELTASTAR);
    freevec(TEMP1);
    freevec(TEMP2);

    return ret_val;
}

/*
 * Indicates whether (1) or not (0) the pair of points from A
 * whose row labels are in pair[] is on opposite sides of the
 * hyperplane by using the signs of the inner products of
```

```
 * each point with the vector (DELTA) thru the origin and
 * perpendicular to the current hyperplane.
 * Note: actually (1/2) is returned if one of the pair is
 * on the hyperplane.
 */

double OppositeSide(MAT *A, int *pair, VEC *DELTA)
{
   VEC *TEMP;
   double temp1, temp2, d;

   TEMP = get_vec(A->n);

   get_row(A,pair[0]-1,TEMP);
   temp1 = in_prod(TEMP,DELTA);
   get_row(A,pair[1]-1,TEMP);
   temp2 = in_prod(TEMP,DELTA);
   d = -(double)(sign(temp1) * sign(temp2) - 1)/2;
   freevec(TEMP);

   return d;
}
```

C.1.2 quadrant.c

```
#include <stdio.h>
#include <math.h>
#include "general.h"

/* Given pairs of vectors (rows of A1 and rows of A2) and estimates
 * of the nuisance location parameters,
 * Quadrant1 returns the value of the interdirection quadrant
 * statistic.
 */

double Quadrant1(MAT *A1, MAT *A2, VEC *THETA1, VEC *THETA2)
{
   double Q;
   int n, adj_n, r1, r2, nhyps1, nhyps2;
   int i, j, z1index, z2index;
   double temp1, temp2, adj1, adj2;
   MAT *T1, *T2;
   MAT *COUNTS1, *COUNTS2;
```

```
    n = A1→m;
    r1 = A1→n;
    r2 = A2→n;
    T1 = get_mat(n,r1);
    T2 = get_mat(n,r2);
    COUNTS1 = get_mat(n,n);
    COUNTS2 = get_mat(n,n);
    cp_mat(A1,T1);
    cp_mat(A2,T2);

    /*
     * This is a patch to adjust the statistic if one of the
     * points happens to be equal to the location estimate
     */

    z1index = Center(T1,THETA1);
    z2index = Center(T2,THETA2);
    adj_n = n − (z1index > 0) − (z2index > 0)
            + (z1index > 0 && z2index > 0);

    nhyps1 = InterdirectionCounts(T1,COUNTS1);
    nhyps2 = InterdirectionCounts(T2,COUNTS2);
    adj1 = (double)((n − r1 + 1) * (n − r1))/(n * (n − 1));
    adj2 = (double)((n − r2 + 1) * (n − r2))/(n * (n − 1));

    Q = 0.0;
    for (i = 1; i ≤ n − 1; ++i) {
        for (j = i + 1; j ≤ n; ++j) {
            temp1 = m_entry(COUNTS1,i−1,j−1)/(adj1 * nhyps1);
            temp2 = m_entry(COUNTS2,i−1,j−1)/(adj2 * nhyps2);
            Q += cos(PI * temp1) * cos(PI * temp2);
        }
    }
    freemat(COUNTS1);
    freemat(COUNTS2);
    freemat(T1);
    freemat(T2);

    return (2 * ((double)(r1 * r2)/adj_n) * Q + (r1 * r2));

}

/* Quadrant2 is a computationally less intensive approximation
 * for Quadrant 1, which must be given additional input for
 * the number of iterations and tuning parameter to be used
```

```
 * in the estimation of covariance matrices.
 */

double Quadrant2(MAT *A1, MAT *A2, VEC *THETA1, VEC *THETA2,
 int iter, double tune)
{

  MAT *SIGMA1, *SIGMA2, *S_INV1, *S_INV2;
  VEC *Y1, *Y2, *X1, *X2;
  VEC *TEMP1, *TEMP2, *TEMP3, *TEMP4;
  int n, r1, r2, i, j;
  double d, d1, d2, s1, s2, t1, t2, n1, n2;

  n = A1->m;
  r1 = A1->n;
  r2 = A2->n;

  SIGMA1 = get_mat(r1,r1);
  SIGMA2 = get_mat(r2,r2);
  S_INV1 = get_mat(r1,r1);
  S_INV2 = get_mat(r2,r2);

  Y1 = get_vec(r1);
  Y2 = get_vec(r1);
  X1 = get_vec(r2);
  X2 = get_vec(r2);

  TEMP1 = get_vec(r1);
  TEMP2 = get_vec(r1);
  TEMP3 = get_vec(r2);
  TEMP4 = get_vec(r2);

  /*
   * Iterative estimation of covariance (M-estimates)
   */

  InitialEstimates(SIGMA1,THETA1,A1);
  InitialEstimates(SIGMA2,THETA2,A2);
  for (i = 0; i < iter; ++i) {
    NextEstimates(SIGMA1,THETA1,A1,tune);
    NextEstimates(SIGMA2,THETA2,A2,tune);
  }

  /*
   * Calculation of statistic
```

```
*/

m_inverse(SIGMA1,S_INV1);
m_inverse(SIGMA2,S_INV2);

d = 0.0;
for(i = 0; i < n - 1; ++i) {

  get_row(A1,i,Y1);
  v_sub(Y1,THETA1,Y1);
  vm_mlt(S_INV1,Y1,TEMP1);
  s1 = sqrt(in_prod(TEMP1,Y1));
  get_row(A2,i,X1);
  v_sub(X1,THETA2,X1);
  vm_mlt(S_INV2,X1,TEMP3);
  t1 = sqrt(in_prod(TEMP3,X1));
  for(j = i + 1; j < n; ++j) {

    get_row(A1,j,Y2);
    v_sub(Y2,THETA1,Y2);
    vm_mlt(S_INV1,Y2,TEMP2);
    n1 = in_prod(TEMP1,Y2);
    s2 = sqrt(in_prod(TEMP2,Y2));
    d1 = (s1 > 0 && s2 > 0 ? n1/(s1 * s2) : 0);
    get_row(A2,j,X2);
    v_sub(X2,THETA2,X2);
    vm_mlt(S_INV2,X2,TEMP4);
    n2 = in_prod(TEMP3,X2);
    t2 = sqrt(in_prod(TEMP4,X2));
    d2 = (t1 > 0 && t2 > 0 ? n2/(t1 * t2) : 0);
    d += d1 * d2;
  }
}
freemat(SIGMA1);
freemat(SIGMA2);
freemat(S_INV1);
freemat(S_INV2);
freevec(X1);
freevec(X2);
freevec(Y1);
freevec(Y2);
freevec(TEMP1);
freevec(TEMP2);
freevec(TEMP3);
freevec(TEMP4);
```

```
  return (2 * (double)(r1 * r2)/n * d + r1 * r2);
}
```

C.1.3 loop.c

```
#include <stdio.h>
#include <math.h>
#include "ranlib.h"
#include "general.h"

#define NOMINAL 0.05

/*
 * This is the main routine which does the repetitions
 * and counts the number of times each of the competing
 * statistics falls in the rejection region.
 */

char *usage = "Usage: %s [p or n] dim N reps delta nu seedfile\n";
char *cmd;

main(int argc, char **argv)
{

  double delta, nu;
  int i, j, dim, r1, r2, N, reps;
  double r, t, u, s, m, n, k, v;
  double Q1, Q2, PS, L, LF, LX, V, VF, VX, U, UF, UX, R, RF;
  double LF_NumDf, LF_DenDf, LX_Df, VF_NumDf, VF_DenDf, VX_Df;
  double UF_NumDf, UF_DenDf, RF_NumDf, RF_DenDf;
  double X, F1, F2, F3, F4;
  VEC *ER, *EI;
  MAT *A1, *A2, *B1, *B2;
  MAT *TEMP1, *TEMP2;
  VEC *THETA1, *THETA2;
  int cLX, cLF, cVX, cVF, cUF, cUX, cRF, cPS, cQ1, cQ2;
  int iLX, iLF, iVX, iVF, iUF, iUX, iRF, iPS, iQ1, iQ2;

  char option;

  long seed1, seed2, g;
  char seedfile[80];
```

```
FILE *seeds;

cmd = *argv;

if (argc < 8) bye();

if (sscanf(argv[1],"%s",&option) ≠ 1) bye();

if (sscanf(argv[2],"%u",&dim) ≠ 1) bye();
if (dim < 1 || dim > MAX_DIM) {
  printf("Invalid value of r\n");
  exit(0);
}

if (sscanf(argv[3],"%u",&N) ≠ 1) bye();
if (N < 2 * r + 1 || N > MAX_N) {
  printf("Invalid value of N\n");
  exit(0);
}

if (sscanf(argv[4],"%u",&reps) ≠ 1) bye();
if (reps < 1) {
  printf("Invalid value of reps\n");
  exit(0);
}

if (sscanf(argv[5],"%lf",&delta) ≠ 1) bye();
if (delta < 0.0 || delta > 0.5) {
  printf("Invalid value of delta\n");
  exit(0);
}

if (sscanf(argv[6],"%lf",&nu) ≠ 1) bye();
if (nu ≤ 0.0) {
  printf("Invalid value of nu\n");
  exit(0);
}

if (sscanf(argv[7],"%s",&seedfile) ≠ 1) bye();

/*
 * Set generator number and current seeds
 */

seeds = fopen(seedfile,"rt");
```

```
while(fscanf(seeds,"%ld %ld %ld",&g,&seed1,&seed2) ≠ EOF);
fclose(seeds);
gscgn(1,&g);
setall(seed1,seed2);

r1 = r2 = dim;
A1 = get_mat(N,r1);
A2 = get_mat(N,r2);
B1 = get_mat(N,r1);
B2 = get_mat(N,r2);
TEMP1 = get_mat(N,r1);
TEMP2 = get_mat(N,r2);
THETA1 = get_vec(r1);
THETA2 = get_vec(r2);
ER = get_vec(r1);
EI = get_vec(r1);

/* Constants needed for statistics */

t = ((r1*r1 + r2*r2 - 5.0 > 0) ?
  pow((r1*r1*r2*r2 - 4.0)/(r1*r1 + r2*r2 - 5.0),0.5) : 1.0);
k = N - (r1 + r2 + 3)/2.0;
u = (r1*r2 - 2)/4.0;
s = min(r1,r2);
m = (abs(r1 - r2) - 1.0)/2.0;
n = (N - r1 - r2 - 2.0)/2.0;
r = max(r1,r2);
v = N - r2 - 1;

LF_NumDf = r1*r2;
LF_DenDf = k*t - 2.0*u;
VF_NumDf = UF_NumDf = s*(2*m + s + 1.0);
VF_DenDf = s*(2*n + s + 1.0);
UF_DenDf = 2*(s*n + 1.0);
RF_NumDf = r;
RF_DenDf = v - r + r2;

X = critchi(NOMINAL,r1*r2);
F1 = critf(NOMINAL,LF_NumDf,LF_DenDf);
F2 = critf(NOMINAL,VF_NumDf,VF_DenDf);
F3 = critf(NOMINAL,UF_NumDf,UF_DenDf);
F4 = critf(NOMINAL,RF_NumDf,RF_DenDf);

cLX = cLF = cVX = cVF = cUF = cUX = cRF = cQ1 = cQ2 = cPS = 0;
```

```
if (option == 'p') {
  printf("\n\n        LF     VF     UF     RF      LX      VX");
  printf("    UX     PS     Q1     Q2\n");
}

for(j = reps; j > 0; --j) {

  /*
   *  Generate the initial random vectors
   */

  rand_dat(A1,nu);
  rand_dat(A2,nu);

  /*
   *  Create dependence via model
   */

  sm_mlt(delta,A1,TEMP1);
  ms_mltadd(TEMP1,A2,(1.0 - delta),B1);
  sm_mlt(delta,A2,TEMP2);
  ms_mltadd(TEMP2,A1,(1.0 - delta),B2);

  /*
   *  Compute statistics
   */

  /* Eigenvalues needed to compute normal theory tests */

  S_Eigenvalues(B1,B2,ER,EI);

  /* Wilks' lambda */

  L = 1.0;
  for (i = 1; i <= r1; ++i)
    L *= (1.0 - v_entry(ER,i-1));

  LF = (pow(L,-1.0/t) - 1.0)*((double)LF_DenDf/LF_NumDf);
  LX = -k*log(L);

  /* Pillai's trace */

  V = 0.0;
  for (i = 1; i <= r1; ++i)
    V += v_entry(ER,i-1);
```

```
VF = V/(s − V)*((double)VF_DenDf/VF_NumDf);
VX = N*V;

/* Hotelling-Lawley trace */

U = 0.0;
for (i = 1; i ≤ r1; ++i)
  U += v_entry(ER,i−1)/(1 − v_entry(ER,i−1));

UF = U/s*((double)UF_DenDf/UF_NumDf);
UX = N*U;

/* Roy's greatest root */

R = v_entry(ER,0);
for (i = 2; i ≤ r1; ++i)
  R = max(R,v_entry(ER,i−1));

RF = R/(1.0 − R)*((double)RF_DenDf/RF_NumDf);

/* Puri-Sen sign statistic */

PS = PuriSen(B1,B2);

/* Quadrant statistics */

OjaMedian(B1,THETA1);
OjaMedian(B2,THETA2);

Q1 = Quadrant1(B1,B2,THETA1,THETA2);
Q2 = Quadrant2(B1,B2,THETA1,THETA2,5,2.0);

/* Reject? */

iLF = (LF > F1);
iVF = (VF > F2);
iUF = (UF > F3);
iRF = (RF > F4);

iLX = (LX > X);
iVX = (VX > X);
iUX = (UX > X);
iPS = (PS > X);
iQ1 = (Q1 > X);
```

```
iQ2 = (Q2 > X);

/*
 *  Update counters
 */

cLF += iLF;
cVF += iVF;
cUF += iUF;
cRF += iRF;
cLX += iLX;
cVX += iVX;
cUX += iUX;
cPS += iPS;
cQ1 += iQ1;
cQ2 += iQ2;

/*
 *   Print value of statistics
 */

if (option == 'p') {
  printf("\n");
  printf("%4d. ",reps - j + 1);

  printf("%6.2f",LF);
  if (iLF) printf("*"); else printf(" ");
  printf("%6.2f",VF);
  if (iVF) printf("*"); else printf(" ");
  printf("%6.2f",UF);
  if (iUF) printf("*"); else printf(" ");
  printf("%6.2f",RF);
  if (iRF) printf("* "); else printf("  ");

  printf("%6.2f",LX);
  if (iLX) printf("*"); else printf(" ");
  printf("%6.2f",VX);
  if (iVX) printf("*"); else printf(" ");
  printf("%6.2f",UX);
  if (iUX) printf("*"); else printf(" ");
  printf("%6.2f",PS);
  if (iPS) printf("*"); else printf(" ");
  printf("%6.2f",Q1);
  if (iQ1) printf("*"); else printf(" ");
  printf("%6.2f",Q2);
```

```
      if (iQ2) printf("*");
    }
  }

  /*
   *  Print empirical powers
   */

  if (option == 'p')
    printf("\n\nEmpirical Powers\n\n");

  printf("%4.2f ",delta);

  printf("%6.4f ",(double)cLF/reps);
  printf("%6.4f ",(double)cVF/reps);
  printf("%6.4f ",(double)cUF/reps);
  printf("%6.4f ",(double)cRF/reps);

  printf("%6.4f ",(double)cLX/reps);
  printf("%6.4f ",(double)cVX/reps);
  printf("%6.4f ",(double)cUX/reps);
  printf("%6.4f ",(double)cPS/reps);
  printf("%6.4f ",(double)cQ1/reps);
  printf("%6.4f\n",(double)cQ2/reps);

  /*
   *  Save current generator and seeds
   */

  gscgn(0,&g);
  getsd(&seed1,&seed2);
  seeds = fopen(seedfile,"at");
  fprintf(seeds,"%ld %ld %ld\n",g,seed1,seed2);
  fclose(seeds);

  return 0;
}

/* exit with message */
int bye()
{
    printf(usage,cmd);
    exit(0);
}
```

C.2 Results in Tabular Form

Following are tables of the simulation results. There are two separate tables on each page dividing the statistics into those which were compared to an F critical value and those which were compared to a chi-square critical value.

Table C.1. $r = 1$, $n = 30$, $\nu = 0.1$, reps = 2500

Δ	**LF**	**VF**	**UF**	**RF**
0.00	0.0688	0.0688	0.0688	0.0688
0.05	0.1544	0.1544	0.1544	0.1544
0.10	0.3836	0.3836	0.3836	0.3836
0.20	0.9380	0.9380	0.9380	0.9380
0.30	0.9928	0.9928	0.9928	0.9928

Table C.2. $r = 1$, $n = 30$, $\nu = 0.1$, reps = 2500

Δ	**LX**	**VX**	**UX**	**PS**	**Q1**	**Q2**
0.00	0.0688	0.0692	0.0772	0.0268	0.0412	0.0412
0.05	0.1544	0.1560	0.1756	0.3188	0.4132	0.4132
0.10	0.3832	0.3892	0.4212	0.6240	0.7284	0.7284
0.20	0.9380	0.9408	0.9488	0.9312	0.9648	0.9648
0.30	0.9928	0.9928	0.9932	0.9940	0.9976	0.9976

Table C.3. $r = 1$, $n = 30$, $\nu = 0.5$, reps = 2500

Δ	LF	VF	UF	RF
0.00	0.0532	0.0532	0.0532	0.0532
0.05	0.0860	0.0860	0.0860	0.0860
0.10	0.2228	0.2228	0.2228	0.2228
0.20	0.7944	0.7944	0.7944	0.7944
0.30	0.9988	0.9988	0.9988	0.9988

Table C.4. $r = 1$, $n = 30$, $\nu = 0.5$, reps = 2500

Δ	LX	VX	UX	PS	Q1	Q2
0.00	0.0532	0.0548	0.0708	0.0228	0.0316	0.0316
0.05	0.0860	0.0884	0.1096	0.0464	0.0612	0.0612
0.10	0.2224	0.2284	0.2748	0.1272	0.1684	0.1684
0.20	0.7944	0.8008	0.8388	0.4672	0.5604	0.5604
0.30	0.9988	0.9988	0.9988	0.8840	0.9276	0.9276

Table C.5. $r = 1$, $n = 30$, $\nu = 1$, reps = 2500

Δ	LF	VF	UF	RF
0.00	0.0500	0.0500	0.0500	0.0500
0.05	0.0928	0.0928	0.0928	0.0928
0.10	0.2280	0.2280	0.2280	0.2280
0.20	0.7728	0.7728	0.7728	0.7728
0.30	0.9996	0.9996	0.9996	0.9996

Table C.6. $r = 1$, $n = 30$, $\nu = 1$, reps = 2500

Δ	**LX**	**VX**	**UX**	**PS**	**Q1**	**Q2**
0.00	0.0500	0.0528	0.0664	0.0320	0.0460	0.0460
0.05	0.0928	0.0964	0.1168	0.0356	0.0564	0.0564
0.10	0.2280	0.2328	0.2708	0.0792	0.1108	0.1108
0.20	0.7728	0.7784	0.8164	0.3052	0.3744	0.3744
0.30	0.9996	0.9996	1.0000	0.7660	0.8308	0.8308

Table C.7. $r = 1$, $n = 30$, $\nu = 10$, reps = 2500

Δ	**LF**	**VF**	**UF**	**RF**
0.00	0.0584	0.0584	0.0584	0.0584
0.05	0.0772	0.0772	0.0772	0.0772
0.10	0.2160	0.2160	0.2160	0.2160
0.20	0.7688	0.7688	0.7688	0.7688
0.30	0.9972	0.9972	0.9972	0.9972

Table C.8. $r = 1$, $n = 30$, $\nu = 10$, reps = 2500

Δ	**LX**	**VX**	**UX**	**PS**	**Q1**	**Q2**
0.00	0.0584	0.0612	0.0764	0.0284	0.0408	0.0408
0.05	0.0772	0.0804	0.0980	0.0264	0.0428	0.0428
0.10	0.2160	0.2220	0.2636	0.0584	0.0776	0.0776
0.20	0.7688	0.7752	0.8076	0.2072	0.2648	0.2648
0.30	0.9972	0.9972	0.9984	0.5920	0.6660	0.6660

Table C.9. $r = 1$, $n = 30$, $df = 1$, reps = 2500

Δ	LF	VF	UF	RF
0.00	0.0652	0.0652	0.0652	0.0652
0.05	0.2508	0.2508	0.2508	0.2508
0.10	0.4788	0.4788	0.4788	0.4788
0.20	0.9448	0.9448	0.9448	0.9448
0.30	0.9956	0.9956	0.9956	0.9956

Table C.10. $r = 1$, $n = 30$, $df = 1$, reps = 2500

Δ	LX	VX	UX	PS	Q1	Q2
0.00	0.0652	0.0672	0.0772	0.0252	0.0392	0.0392
0.05	0.2508	0.2536	0.2768	0.0944	0.1288	0.1288
0.10	0.4784	0.4840	0.5200	0.2516	0.3128	0.3128
0.20	0.9448	0.9460	0.9556	0.6944	0.7700	0.7700
0.30	0.9956	0.9960	0.9964	0.9632	0.9788	0.9788

Table C.11. $r = 1$, $n = 30$, $df = 5$, reps = 2500

Δ	LF	VF	UF	RF
0.00	0.0524	0.0524	0.0524	0.0524
0.05	0.0800	0.0800	0.0800	0.0800
0.10	0.2208	0.2208	0.2208	0.2208
0.20	0.7804	0.7804	0.7804	0.7804
0.30	0.9984	0.9984	0.9984	0.9984

Table C.12. $r = 1$, $n = 30$, $df = 5$, reps = 2500

Δ	**LX**	**VX**	**UX**	**PS**	**Q1**	**Q2**
0.00	0.0524	0.0540	0.0640	0.0244	0.0384	0.0384
0.05	0.0800	0.0828	0.1068	0.0296	0.0428	0.0428
0.10	0.2208	0.2268	0.2612	0.0852	0.1116	0.1116
0.20	0.7804	0.7848	0.8212	0.3588	0.4368	0.4368
0.30	0.9984	0.9984	0.9992	0.8224	0.8860	0.8860

Table C.13. $r = 2$, $n = 30$, $\nu = 0.1$, reps = 2500

Δ	**LF**	**VF**	**UF**	**RF**
0.00	0.1152	0.1072	0.1196	0.1960
0.05	0.1528	0.1420	0.1588	0.2740
0.10	0.3672	0.3620	0.3696	0.5148
0.20	0.9704	0.9708	0.9700	0.9916
0.30	1.0000	1.0000	1.0000	1.0000

Table C.14. $r = 2$, $n = 30$, $\nu = 0.1$, reps = 2500

Δ	**LX**	**VX**	**UX**	**PS**	**Q1**	**Q2**
0.00	0.1156	0.1088	0.1620	0.0852	0.0628	0.0576
0.05	0.1528	0.1444	0.2240	0.2980	0.3364	0.3208
0.10	0.3672	0.3648	0.4912	0.6304	0.7564	0.7356
0.20	0.9704	0.9708	0.9916	0.9696	0.9968	0.9948
0.30	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table C.15. $r = 2$, $n = 30$, $\nu = 0.5$, reps = 2500

Δ	**LF**	**VF**	**UF**	**RF**
0.00	0.0516	0.0452	0.0560	0.1652
0.05	0.0860	0.0788	0.0936	0.2188
0.10	0.2084	0.2016	0.2124	0.3952
0.20	0.8784	0.8788	0.8728	0.9404
0.30	1.0000	1.0000	1.0000	1.0000

Table C.16. $r = 2$, $n = 30$, $\nu = 0.5$, reps = 2500

Δ	**LX**	**VX**	**UX**	**PS**	**Q1**	**Q2**
0.00	0.0516	0.0456	0.1136	0.0828	0.0604	0.0504
0.05	0.0860	0.0788	0.1720	0.1004	0.0844	0.0760
0.10	0.2088	0.2024	0.3296	0.1992	0.2096	0.1984
0.20	0.8784	0.8792	0.9408	0.6480	0.7960	0.7868
0.30	1.0000	1.0000	1.0000	0.9828	0.9984	0.9988

Table C.17. $r = 2$, $n = 30$, $\nu = 1$, reps = 2500

Δ	**LF**	**VF**	**UF**	**RF**
0.00	0.0488	0.0452	0.0516	0.1548
0.05	0.0772	0.0696	0.0824	0.2220
0.10	0.2268	0.2088	0.2316	0.4108
0.20	0.8648	0.8632	0.8644	0.9232
0.30	0.9996	0.9996	0.9996	1.0000

Table C.18. $r = 2$, $n = 30$, $\nu = 1$, reps = 2500

Δ	**LX**	**VX**	**UX**	**PS**	**Q1**	**Q2**
0.00	0.0488	0.0456	0.1124	0.0844	0.0556	0.0456
0.05	0.0780	0.0704	0.1692	0.0872	0.0712	0.0644
0.10	0.2268	0.2120	0.3560	0.1540	0.1748	0.1536
0.20	0.8648	0.8636	0.9256	0.5100	0.6320	0.6208
0.30	0.9996	0.9996	1.0000	0.9300	0.9876	0.9892

Table C.19. $r = 2$, $n = 30$, $\nu = 10$, reps = 2500

Δ	**LF**	**VF**	**UF**	**RF**
0.00	0.0480	0.0416	0.0524	0.1552
0.05	0.0740	0.0680	0.0788	0.2116
0.10	0.2124	0.2048	0.2200	0.4040
0.20	0.8700	0.8680	0.8672	0.9312
0.30	1.0000	1.0000	1.0000	1.0000

Table C.20. $r = 2$, $n = 30$, $\nu = 10$, reps = 2500

Δ	**LX**	**VX**	**UX**	**PS**	**Q1**	**Q2**
0.00	0.0480	0.0436	0.1144	0.0876	0.0540	0.0504
0.05	0.0740	0.0688	0.1608	0.0968	0.0756	0.0616
0.10	0.2124	0.2052	0.3476	0.1480	0.1376	0.1260
0.20	0.8704	0.8712	0.9336	0.4048	0.5004	0.4900
0.30	1.0000	1.0000	1.0000	0.8664	0.9640	0.9672

Table C.21. $r = 2$, $n = 30$, $df = 1$, reps = 2500

Δ	LF	VF	UF	RF
0.00	0.1136	0.1076	0.1168	0.1700
0.05	0.3628	0.3556	0.3664	0.4728
0.10	0.6192	0.6172	0.6208	0.7132
0.20	0.9900	0.9900	0.9904	0.9968
0.30	1.0000	1.0000	1.0000	1.0000

Table C.22. $r = 2$, $n = 30$, $df = 1$, reps = 2500

Δ	LX	VX	UX	PS	Q1	Q2
0.00	0.1140	0.1076	0.1492	0.0748	0.0500	0.0416
0.05	0.3628	0.3564	0.4360	0.1892	0.2108	0.1860
0.10	0.6192	0.6188	0.7000	0.4340	0.5280	0.4944
0.20	0.9900	0.9900	0.9964	0.9160	0.9740	0.9680
0.30	1.0000	1.0000	1.0000	0.9984	1.0000	0.9996

Table C.23. $r = 2$, $n = 30$, $df = 5$, reps = 2500

Δ	LF	VF	UF	RF
0.00	0.0460	0.0436	0.0508	0.1564
0.05	0.0852	0.0796	0.0900	0.2208
0.10	0.2220	0.2100	0.2252	0.4152
0.20	0.8896	0.8900	0.8884	0.9432
0.30	1.0000	1.0000	1.0000	1.0000

Table C.24. $r = 2$, $n = 30$, $df = 5$, reps = 2500

Δ	**LX**	**VX**	**UX**	**PS**	**Q1**	**Q2**
0.00	0.0460	0.0440	0.1064	0.0804	0.0492	0.0448
0.05	0.0852	0.0800	0.1668	0.1012	0.0944	0.0840
0.10	0.2220	0.2104	0.3592	0.1856	0.1796	0.1664
0.20	0.8896	0.8920	0.9512	0.6028	0.7552	0.7484
0.30	1.0000	1.0000	1.0000	0.9696	0.9980	0.9984

Table C.25. $r = 3$, $n = 30$, $\nu = 0.1$, reps = 1000

Δ	**LF**	**VF**	**UF**	**RF**
0.00	0.135	0.116	0.149	0.373
0.05	0.185	0.162	0.201	0.466
0.10	0.375	0.357	0.385	0.692
0.00	0.135	0.116	0.149	0.373
0.05	0.185	0.162	0.201	0.466

Table C.26. $r = 3$, $n = 30$, $\nu = 0.1$, reps = 1000

Δ	**LX**	**VX**	**UX**	**Q1**	**Q2**
0.00	0.136	0.116	0.257	0.074	0.058
0.05	0.189	0.160	0.355	0.279	0.223
0.10	0.375	0.354	0.601	0.754	0.697
0.00	0.136	0.116	0.257	0.074	0.058
0.05	0.189	0.160	0.355	0.279	0.223

Table C.27. $r = 3$, $n = 30$, $\nu = 0.5$, reps = 1000

Δ	**LF**	**VF**	**UF**	**RF**
0.00	0.057	0.039	0.067	0.356
0.05	0.088	0.073	0.105	0.404
0.10	0.228	0.205	0.245	0.620
0.00	0.057	0.039	0.067	0.356
0.05	0.088	0.073	0.105	0.404

Table C.28. $r = 3$, $n = 30$, $\nu = 0.5$, reps = 1000

Δ	**LX**	**VX**	**UX**	**Q1**	**Q2**
0.00	0.057	0.039	0.195	0.059	0.037
0.05	0.090	0.073	0.264	0.106	0.088
0.10	0.230	0.205	0.495	0.262	0.216
0.00	0.057	0.039	0.195	0.059	0.037
0.05	0.090	0.073	0.264	0.106	0.088

Table C.29. $r = 3$, $n = 30$, $\nu = 1$, reps = 1000

Δ	**LF**	**VF**	**UF**	**RF**
0.00	0.052	0.038	0.059	0.358
0.05	0.079	0.063	0.096	0.422
0.10	0.212	0.200	0.234	0.601
0.00	0.052	0.038	0.059	0.358
0.05	0.079	0.063	0.096	0.422

Table C.30. $r = 3$, $n = 30$, $\nu = 1$, reps = 1000

Δ	**LX**	**VX**	**UX**	**Q1**	**Q2**
0.00	0.053	0.038	0.200	0.075	0.060
0.05	0.080	0.063	0.271	0.084	0.063
0.10	0.217	0.198	0.490	0.190	0.145
0.00	0.053	0.038	0.200	0.075	0.060
0.05	0.080	0.063	0.271	0.084	0.063

REFERENCES

Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis* (second edition). Wiley, New York.

Apostol, T. M. (1957). *Mathematical Analysis: A Modern Approach to Advanced Calculus.* Addison-Wesley, Reading, Massachusetts.

Barrodale, I., & Roberts, F. D. K. (1974). Solution of an overdetermined system of equations in the L_1-norm. *Communications of the Association of Computing Machinery, 17*(6), 319–320.

Bartlett, M. S. (1938). Further aspects of the theory of multiple regression. *Proceedings of the Cambridge Philosophical Society, 34*, 33–40.

Becker, R. A., Chambers, J. M., & Wilks, A. R. (1988). *The New S Language: A Programming Environment for Data Analysis and Graphics.* Wadsworth & Brooks/Cole, Pacific Grove, California.

Behnke, M., Eyler, F. D., Conlon, M., Woods, N. S., & Thomas, V. J. (1993). The relationship between umbilical cord and infant blood gases and developmental outcome in very low birth weight infants. *Clinical Obstetrics and Gynecology, 36*(1), 73–81.

Bhuchongkul, S. (1964). A class of nonparametric tests for independence in bivariate populations. *Annals of Mathematical Statistics, 35*, 138–149.

Blomqvist, N. (1950). On a measure of dependence between two random variables. *Annals of Mathematical Statistics, 21*, 593–600.

Box, G. E. P. (1949). A general distribution theory for a class of likelihood criteria. *Biometrika, 36*, 317–346.

Daniels, H. E. (1944). Measures of correlation in the universe of sample permutations. *Biometrika, 33*, 129–135.

de Wet, T., & Randles, R. H. (1987). On the effect of substituting parameter estimators in limiting χ^2 U- and V-statistics. *Annals of Statistics, 15*, 398–412.

Dielman, T. E. (1992). Computational algorithms for least absolute value regression. In Dodge, Y. (Ed.), *L_1-Statistical Analysis and Related Methods.* Elsevier Science Publishers B.V., The Netherlands.

Farlie, D. J. G. (1960). The performance of some correlation coefficients for a general bivariate distribution. *Biometrika, 47*, 307–323.

Farlie, D. J. G. (1961). The asymptotic efficiency of Daniels' generalized correlation coefficients. *Journal of the Royal Statistical Society, Series B*, *23*, 128–142.

Galton, F. (1888). Correlations and their measurement. *Proceedings of the Royal Statistical Society*, *45*, 135.

Graybill, F. A. (1976). *Theory and Application of the Linear Model.* Duxbury Press, North Scituate, Massachusetts.

Graybill, F. A. (1983). *Matrices with Applications in Statistics* (second edition). Wadsworth International Group, Belmont, California.

Greiner, R. (1909). Über das fehlersystem der kollektivmasslehre. *Zeitschrift für Mathematik und Physik*, *57*, 121,225,337.

Hájek, J., & Šidák, Z. (1967). *Theory of Rank Tests.* Academic Press, New York.

Hannan, E. J. (1956). The asymptotic powers of certain tests based on multiple correlations. *Journal of the Royal Statistical Society, Series B*, *18*, 227–233.

Hettmansperger, T. P., Nyblom, J., & Oja, H. (1992). On multivariate notions of sign and rank. In Dodge, Y. (Ed.), *L_1-Statistical Analysis and Related Methods.* Elsevier Science Publishers B.V., The Netherlands.

Hoeffding, W. (1948). A non-parametric test of independence. *Annals of Mathematical Statistics*, *23*, 169.

Hotelling, H. (1936). Relations between two sets of variates. *Biometrika*, *28*, 321–377.

Hotelling, H. (1947). Selected techniques of statistical analysis. In Eisenhart, C. (Ed.), *Multivariate Quality Control Illustrated by the Air Testing of Sample Bombsights.* McGraw-Hill, New York.

Hotelling, H. (1951). A generalized T test and measure of multivariate dispersion. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, *47*, 23–41.

Iverson, H. K. (1982). *Asymptotic Properties of U-Statistics with Estimated Parameters.* Ph.D. dissertation, University of Iowa, Iowa City.

Iverson, H. K., & Randles, R. H. (1989). The effects on convergence of substituting parameter estimates into U-statistics and other families of statistics. *Probability Theory and Related Fields*, *81*, 453–471.

Jan, S. (1991). *Interdirection Tests for Repeated Measures and One-Sample Multivariate Location Problems.* Ph.D. dissertation, University of Florida, Gainesville.

Johnson, M. E. (1987). *Multivariate Statistical Simulation.* Wiley, New York.

Johnson, R. A., & Wichern, D. W. (1988). *Applied Multivariate Statistical Analysis* (second edition). Prentice-Hall, Englewood Cliffs, New Jersey.

Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika, 30*, 81.

Kendall, M. G. (1970). *Rank Correlation Methods* (fourth edition). Charles Griffin & Company Limited, London.

Kernighan, B. W., & Ritchie, D. M. (1988). *The C Programming Language* (second edition). Prentice-Hall, Englewood Cliffs, New Jersey.

Knuth, D. E. (1970). *The TEXbook.* Addison-Wesley Publishing Company, Reading, Massachusetts.

Konijn, H. S. (1954). On the power of certain tests for independence in bivariate populations. *Annals of Mathematical Statistics*, *25*, 300–323.

Kruskal, W. H. (1958). Ordinal measures of association. *Journal of the American Statistical Association*, *53*(284), 814–861.

Lamport, L. (1986). LATEX*: A Document Preparation System.* Addison-Wesley Publishing Company, Reading, Massachusetts.

Lawley, D. N. (1938). A generalization of Fisher's z-test. *Biometrika*, *30*, 180–187.

Lee, A. J. (1990). *U-statistics: Theory and Practice.* M. Dekker, New York.

Lee, Y. (1971). Distribution of the canonical correlations and asymptotic expansions for distributions of certain independence test statistics. *The Annals of Mathematical Statistics*, *42*, 526–537.

Merchant, J. A., et al. (1975). Responses to cotton dust. *Archives of Environmental Health*, *30*, 222–229.

Miller, K. S. (1964). *Multidimensional Gaussian Distributions.* Wiley, New York.

Miller, K. S. (1968). Some multivariate t-distributions. *The Annals of Mathematical Statistics*, *39*, 1605–1609.

Morrison, D. F. (1976). *Multivariate Statistical Methods* (second edition). McGraw-Hill Book Company, New York.

Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory.* Wiley, New York.

Niinimaa, A. O. (1992). The calculation of the Oja multivariate median. In Dodge, Y. (Ed.), *L_1-Statistical Analysis and Related Methods.* Elsivier Science Publishers B.V., The Netherlands.

Oja, H. (1983). Descriptive statistics for multivariate distributions. *Statistics and Probability Letters*, *1*(6), 327–332.

Oja, H., & Niinimaa, A. (1985). Asymptotic properties of the generalized median in the case of multivariate normality. *Journal of the Royal Statistical Society, Series B*, *47*, 372–377.

Pearson, K. (1896). Mathematical contributions to the theory of evolution—iii. Regression, Heredity and Panmixia. *Philosophical Transactions of the Royal Society of London, Series A*, *187*, 253–318.

Pearson, K. (1920). Notes on the history of correlation. *Biometrika*, *13*, 25.

Peters, D., & Randles, R. H. (1990). A multivariate signed-rank test for the one-sample location problem. *Journal of the American Statistical Association*, *85*, 552–557.

Puri, M. L., & Sen, P. K. (1971). *Nonparametric Methods in Multivariate Analysis.* Wiley, New York.

Randles, R. H. (1982). On the asymptotic normality of statistics with estimated parameters. *Annals of Statistics*, *10*, 462–474.

Randles, R. H. (1989). A distribution-free multivariate sign test based on interdirections. *Journal of the American Statistical Association*, *84*, 1045–1050.

Randles, R. H., Broffitt, J. D., Ramberg, J. S., & Hogg, R. V. (1978). Generalized linear and quadratic discriminant functions using robust estimates. *Journal of the American Statistical Association*, *73*, 564–568.

Randles, R. H., & Wolfe, D. A. (1979). *Introduction to the Theory of Nonparametric Statistics.* Wiley, New York.

SAS Institute Inc., (1989). *SAS/STAT User's Guide, Volume 1* (fourth edition). Cary, North Carolina.

Searle, S. R. (1982). *Matrix Algebra Useful for Statistics.* Wiley, New York.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics.* Wiley, New York.

Sinha, B. K., & Wieand, H. S. (1977). Multivariate nonparametric tests for independence. *Journal of Multivariate Analysis*, *7*, 572–583.

Spearman, C. (1904). The proof and measurement of association between two things. *American Journal of Psychology*, *15*, 72–101.

Stuart, A. (1954). The correlation between variate values and ranks in samples from a continuous distribution. *British Journal of Statistical Psychology*, *7*, 37–44.

Wilks, S. S. (1935). On the independence of k sets of normally distributed statistical variables. *Econometrica*, *3*, 309–326.

Wolfram, S. (1988). *Mathematica: A System for Doing Mathematics by Computer.* Addison-Wesley Publishing Company, Redwood City, California.

Yule, G. U., & Kendall, M. G. (1940). *An Introduction to the Theory of Statistics.* Charles Griffin & Company Limited, London.

BIOGRAPHICAL SKETCH

I was born Peter William Trust in the small town of Shell Lake, Wisconsin, on 9 May 1965. My early childhood was not without hardship, yet I was fortunate enough to not ever be aware of it, save for the death of my father, William, when I was around six years old. After moving to Minnesota, my mother, Jackie, had the incredibly good fortune to meet a man who has fulfilled every definition of the word dad. Harold Gieser married my mother on 5 March 1973 and adopted me and my younger sister, Wendy, shortly thereafter. With his children, Kathy and Greg, from a previous marriage and the birth of Angi in 1975, the Gieser family was complete.

My interest in scholastics did not really show itself until late in my high school years, when I was lucky enough to get involved with a group of friends who valued education. After graduating from Park Rapids Area High School in Minnesota in 1983, I started college at St. Johns University with the hope of completing a degree in mathematics. Alas, it was not meant to be. After running short on funds, I transferred to St. Cloud State University to finish my final two years. It was at this time that a significant turning point in my academic career occurred. After taking several statistics courses to fulfill a math requirement, I knew that that was what I wanted to do. After graduating from SCSU in 1988 with a B.A. in mathematics and a minor in statistics, I traveled to Gainesville, Florida, to begin the quest for a Ph.D. in statistics. Five years later that quest has been achieved.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Ronald H. Randles

Ronald H. Randles, Chairman
Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Jane F. Pendergast

Jane F. Pendergast
Research Associate Professor

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

P. V. Rao

Pejaver V. Rao
Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Dennis D. Wackerly

Dennis D. Wackerly
Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

James J. Algina

James J. Algina
Professor of Foundations of Education

This dissertation was submitted to the Graduate Faculty of the Department of Statistics in the College of Liberal Arts and Sciences and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

December 1993

____
Dean, Graduate School